Approximate Variational Coupled Cluster Theories

- A thesis submitted for the degree of Doctor of Philosophy -

- by -

James B. Robinson

Cardiff University

July 2012

Abstract

This thesis presents a new family of single-determinantal *ab initio* electronic structure models ideal for the black-box computational investigation of molecular phenomena. These theories mimic Variational Coupled Cluster and achieve an approximately fulfilled variational upper bound property on the exact ground-state Schrödinger energy eigenvalue, while not exceeding the limiting computational complexity of the well-known CCSD method, and without sacrificing any other advantageous methodological property. In particular, these *Approximate Variational Coupled Cluster Theories* are formulated through the minimization of a rigorously extensive and orbitally-invariant functional that treats certain limiting systems exactly. Unlike CCSD and related methods, for which it is extremely problematic, these theories are highly robust to the breakdown of the Hartree-Fock approximation that occurs when the single Slater determinant of minimal energy becomes an inadequate qualitative description of the true electronic wavefunction.

Furthermore, presented results suggest that when the essential physics of strong non-dynamic correlation is captured by a near-variational ansatz, remaining dynamic correlation effects may be legitimately included perturbatively, with implications for the design of future electronic structure models because the failure of methods such as CCSD(T) to describe the dissociation of multiple bonds may be ascribed to the inadequate non-variational description of the electronic structure at the CCSD level, and not to a breakdown of perturbation theory; Optimized-Orbital Quasi-Variational Coupled Cluster Doubles with a perturbative treatment of triple excitations (OQVCCD(T)) is capable of predicting a physically correct and quantitatively accurate potential energy curve for diatomic nitrogen, N_2, which has not been achieved, at the time of writing, by any other practical ($\mathcal{O}(o^2v^4)$-iterative $\mathcal{O}(o^3v^4)$-non-iterative) method based on Restricted Hartree-Fock theory. The method is demonstrated to be additionally suitable for the black-box description of singlet multiradicals through application to model hydrogen chains.

Contents

Chapter 1

Motivation

A central theoretical and computational problem in the fields of quantum chemistry and molecular physics is the purely first principles (or *ab initio*) prediction of molecular electronic structure, because it is the electronic structure of a molecule that primarily determines its physical and chemical properties. *Ab initio* electronic structure calculations attempt to find an approximate solution to the Born-Oppenheimer electronic Schrödinger equation, and have, in recent years, become indispensable tools for the investigation of molecular phenomena. The construction of new algorithms for the more efficient or more accurate computational calculation of molecular electronic structure therefore continues to be a very active area of research.

The standard and widely accepted approach to the treatment of this problem is to first make the Hartree-Fock[1, 2] (HF) approximation, which constructs the single-determinantal reference wavefunction of minimal energy, solving the resulting equations by a Self-Consistent Field procedure, and assumes this wavefunction to be an adequate description of the ground-state electronic structure. This approximation corresponds to a mean-field treatment of the Coulomb and exchange interactions between electrons.

The Hartree-Fock approximation typically accounts for almost all of the molecular energy, but the small remaining energy, the 'correlation energy', is extremely important for the correct description of phenomena such as chemical bonding.

Electron correlation may be further conceptually divided into 'dynamic' and 'non-dynamic' (or static) correlation. These terms refer respectively to the energy associated with the instantaneous correlated motion of the electrons that is not captured by the HF mean-field treatment, and to the character of the system that cannot be captured easily by the underlying single-determinantal reference wavefunction approximation. One of the key challenges in obtaining sufficient accuracy for reliable and quantitative prediction of empirical phenomena is to obtain the energy associated with the correlated motion of the electrons sufficiently completely. Corrections to Hartree-Fock theory therefore must be computed by some scheme.

Low-order corrections may be determined from some form of many-body perturbation theory, such as the Møller-Plesset method[3], which partitions the electronic Hamiltonian into the Hamiltonian for which the Hartree-Fock reference wavefunction is the exact eigenfunction, and the Fluctuation Potential, which describes the difference between the Hartree-Fock Hamiltonian and the true electronic Hamiltonian, with the HF Slater determinant playing the role of the zeroth-order wavefunction. However, the convergence of the perturbation series is often a troubling consideration[4–6] and motivates the use of other theoretical methodologies, although these are often used in conjunction with perturbation theory for reasons of cost.

The true electronic wavefunction can be expanded exactly in the basis of the determinants that can be generated by the replacement of one or more of the orbitals occupied in the reference wavefunction with unoccupied orbitals. The conceptually (although not computationally) simplest alternative to perturbation theory that exploits this expansion is Configuration Interaction[7] (CI), in which the exact wavefunction is simply expanded linearly in the determinantal N-electron basis. When all possible determinants are included it is referred to as Full Configuration Interaction (FCI). Unfortunately, although the determinantal series is finite, at least in a finite 1-electron basis set, the series still typically must be truncated for computational practicality. An attempt is therefore made to capture a representative subset of the terms in the determinantal series by truncation to, for example, at most twofold excitations of the reference determi-

nant. In the case of CI, this defines the CISD method. However, CISD is not extensive, meaning that calculated energies do not scale physically with the size of the system under consideration, which limits the applicability of the scheme.

Early extensive or approximately extensive methods attempted to eliminate unphysical unlinked cluster contributions, and included, for example, the Coupled Pair Functional[8, 9] (CPF), the Coupled Electron Pair Approximation[10–15] (CEPA), the Linear Coupled Pair Many-Electron Theory[15, 16] (LCPMET), also known as CEPA(0), and the very simple Davidson correction to CISD[17]. However, many of these approaches introduced other problems, such as discrepancies from CISD/FCI for the simplest molecules involving the correlation of just 2 electrons, and a lack of invariance to rotations of the underlying orbital spaces. However, more recent work[18] has resulted in more sophisticated, orbitally invariant approaches.

The Traditional Coupled Cluster[19–26] (TCC) method has emerged as the standard approach (based on a single-determinantal reference wavefunction) to the treatment of the electron correlation problem, sometimes combined with perturbative approximations. The ansatz of Coupled Cluster (CC) theory is that the exact wavefunction should be parameterised exponentially in the determinantal basis, and not linearly, as it is in CI theory. The usual approach is to model the exact wavefunction as the action of an exponential 'cluster' operator on the Hartree-Fock reference wavefunction, then, in the case of TCC, to project the Schrödinger equation onto the appropriate manifold of excited determinants and iteratively solve the resulting equations to determine the cluster operator, and thus the molecular energy. TCC calculations are, again for reasons of computational cost, most commonly performed with a cluster operator that generates only single and double excitations of the reference wavefunction, and this is termed CCSD. The TCC energy is rigorously extensive, even for a truncated cluster operator, and exactly equivalent to the FCI energy when the cluster operator is complete. In addition, the Campbell-Baker-Hausdorff expansion[27] of the similarity-transformed Hamiltonian naturally truncates at the fourth power of the cluster operators, ensuring that the computational effort of a TCC calculation is always of polynomial, as opposed to factorial, complexity.

When the underlying single-determinantal reference wavefunction of Hartree-Fock theory is a good approximation to the ground-state molecular electronic structure, CCSD performs extremely well. However, some aspects of the TCC formalism are problematic; the similarity-transformed Hamiltonian is not Hermitian, which complicates the calculation of both ground- and excited-state properties. In addition, the non-Hermitian nature of this effective Hamiltonian, combined with the projective determination of the equations to be solved means that, unlike methods such as Variational CI, a calculated TCC energy is not an upper bound on the exact ground-state Schrödinger energy eigenvalue. In contrast, fully variational methods have the advantage that the error in the minimized energy is second order in remaining errors in the wavefunction or parameterisation. Historically, this property led to the preferred use of Variational CI to account for electron correlation, but, in recent decades, the importance of the guaranteed energy extensivity of TCC has led to a decline in the use of CI.

Furthermore, situations exist in which the Hartree-Fock approximation breaks down, such that the single-determinantal reference wavefunction becomes a poor approximation to the true ground-state wavefunction. In these circumstances, for which non-dynamic correlation is said to be strong, there are questions as to whether, for a finite truncation of the cluster operator, TCC can describe phenomena such as dissociating molecules or excited states faithfully. In practice, CCSD and related methods can fail catastrophically, especially when applied to difficult problems such as the dissociation of multiply-bonded molecules. These failures often emerge as the prediction of unphysical maxima in potential energy surfaces, followed by non-variational collapse to energies significantly below FCI[28]. In order to circumvent these failures within the methodology of TCC theory, one must compute increasingly higher-order corrections through the explicit inclusion of triple and higher excitations. This comes with a steep computational cost, however; the CCSDT method scales with a computational complexity of $\mathcal{O}(N^8)$, where N is some measure of the system size, such as the number of electrons. This is in contrast to the more satisfactory $\mathcal{O}(N^6)$ complexity of CCSD. The CCSDTQ method is even more computationally demanding, with $\mathcal{O}(N^{10})$ complexity.

This problem can be potentially resolved by adopting a method that employs a reference wavefunction consisting of multiple determinants, for example, multireference CI (MRCI)[29, 30], and related formulations that correct approximately for extensivity errors[31–33]. However, unlike for methods based on HF theory, for which the reference wavefunction is trivially determined as the single Slater determinant of minimal energy, the determinants to be included in the multi-determinantal reference wavefunction must be chosen manually, and chosen well, and these methods are thus difficult to use in a black-box fashion, especially on large molecules. They are also often highly expensive in terms of computational effort, and encounter problems due to the lack of rigorous extensivity of the energy. Other, more novel approaches to the treatment of non-dynamic correlation have been proposed, such as the active-space CC methods of Head-Gordon[34–36], and the spin-flip[37, 38] and double-ionization-potential[39, 40] EOM methods, but an all-purpose method has yet to emerge. Thus, there exist systems for which practical single-reference methods such as CCSD yield inadequate descriptions of the electronic structure, and also for which multireference methods are impractical. A method that adequately describes phenomena in this niche would be highly desirable.

Numerous studies have shown[28, 41–45] that the poor performance of CCSD in the regime of strong non-dynamic correlation is not necessarily the fault of the Coupled Cluster wavefunction ansatz, but, in fact, arises from the projective determination of the cluster amplitudes inherent to the TCC approach. Alternative Coupled Cluster methods, such as Unitary Coupled Cluster[46–50] (UCC), and Variational Coupled Cluster[51] (VCC), have been proposed. In addition, asymmetric expectation value expressions exist, such as Improved Coupled Cluster[52] (ICC) and Extended Coupled Cluster[53] (ECC), which suggest hierarchies of methods stepping systematically from TCC to VCC depending on the level of the truncation of the series, although they possess no rigorous guarantee of extensivity at arbitrary truncation, except at the extremes of the hierarchies. These studies have confirmed the alternative Coupled Cluster ansätze to be significantly more robust to the breakdown of the Hartree-Fock approximation than TCC. The excellent performance of the VCC method, in particular, for which the cluster

operator is optimized by variational minimization of the quantum-mechanical energy expectation value rather than by projection, can be attributed to its rigorous upper bound property. This means that it is mathematically impossible for VCC-calculated energies to lie below the corresponding ground-state Schroödinger energy eigenvalue. In practice, this means that VCC potential energy curves do not predict the unphysical maxima and subsequent catastrophic failure that is problematic for TCC. In fact, there is also an asymmetry between VCC and TCC in the levels of excited determinants required to achieve a physically correct treatment of difficult problems such as multiple bond breaking, with VCC requiring only single and double excitations and TCC requiring a full treatment of singles, doubles, triples and quadruples[28]. This asymmetry can be understood in light of the analyses of Kutzelnigg[13, 54].

Unfortunately, each of the alternative CC methods scale unfavourably in computational complexity relative to TCC. For the VCC method, in particular, there is no analogue of the Campbell-Baker-Hausdorff expansion, and the computational difficulty of the VCC method therefore scales factorially with system size, making it impractical to apply to all but the simplest of molecular systems. Even the Quadratic Coupled Cluster[55] (QCC) method, which represents the simplest possible correction of TCC towards VCC, is significantly more computationally demanding, and has therefore not been widely adopted. In fact, the recent study of Evangelista[56] confirmed that the additional accuracy of these methods is usually significantly outweighed by their increased computational cost.

This thesis investigates whether there exist tractable approximations to the VCC method with computational complexities not exceeding that of TCC, and whether these approximations can remain true to the spirit of Coupled Cluster theory and preserve its notable methodological properties, such as rigorous extensivity and equivalence to FCI for a complete cluster operator. It is further investigated whether these single-reference *Approximate Variational Coupled Cluster Theories* are robust enough to describe problems such as multiple bond breaking, and other situations involving strong non-dynamic electron correlation for which TCC fails to perform adequately.

Chapter 2

Electronic Structure Theory

The chemical properties of the two simplest chemical elements could not be more different; hydrogen (H) is highly reactive, whereas helium (He) is almost completely inert. These differences can be explained by the electronic structures of the constituent atoms. The electrons present in a helium atom are paired and stable, but it is energetically favourable for a hydrogen atom to seek to pair its single electron by, for example, forming H_2. The electronic structure is therefore the essential property determining the chemical reactivity and physical properties of a substance. In fact, the electronic structure of a molecule may be determined entirely from the laws of physics; it was the physicist Paul Dirac who said[57],

> *The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.*

The field of *ab initio* quantum chemistry concerns this application of the laws of physics, and quantum mechanics in particular, to the study of chemistry. Dirac further continued,

> *It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.*

There has been much success in the design and application of these methods, and the use of computational tools to predict chemical phenomena has become increasingly widespread over the last few decades. These techniques are now routinely applied across diverse fields, such as physics, chemistry, biology and engineering, and for tasks from the rationalization of observed phenomena and the elucidation of mechanisms inaccessible to empirical study to the prediction of the properties of hypothetical molecular designs.

Despite these successes, however, open problems remain. In particular, this thesis explores the hypothesis that it is possible to construct practical quantum-chemical methods robust enough to yield quantitatively valid results even when the underlying approximation scheme on which they are based breaks down. If true, this would open a new class of systems to a black-box treatment of molecular electronic structure. This chapter presents a summary of contemporary electronic structure theory, and introduces this problem.

2.1 Molecular Quantum Mechanics

The predominant Classical view was that matter, such as atoms and molecules, consisted of particles, whereas light consisted of electromagnetic waves. As the 20^{th} century dawned however, new evidence, such as the famous photoelectric experiment, for which Einstein won the Nobel Prize, confirmed that light could behave like particles (now called 'photons'), and that particles could display wave properties, for example due to the observation of electron diffraction.

It is now well understood and accepted that there is a duality between particles and waves captured by the famous de Broglie relationship between wavelength, λ, (a wave property) and momentum, p, (a particle property),

$$\lambda = \frac{h}{p}, \tag{2.1}$$

where h is the Planck constant. Objects of everyday proportions are typically well treated by classical mechanics because their de Broglie wavelengths are many orders of magnitude smaller than the particle itself.

When it comes to chemistry, however, classical mechanics fails completely to correctly model the true physics between or within atoms. This is because the de Broglie wavelength of the electron, in particular, is on the order of atomic length scales, and a wave picture of electrons in atoms becomes essential. This wave theory of matter is called Quantum Mechanics, and is summarized by a set of axioms given in, for example, Reference [58]. The fundamental principle is that the state of a system is described by a mathematical object called a wavefunction, Ψ, that satisfies the well-known Schrödinger Equation,

$$\hat{H}\Psi(\mathbf{x},t) = i\frac{\partial\Psi(\mathbf{x},t)}{\partial t}, \tag{2.2}$$

where atomic units[59] are used. If the wavefunction describes a steady-state phenomenon, then it becomes separable in spatial and temporal coordinates,

$$\Psi(\mathbf{x},t) = \psi(\mathbf{x})\phi(t), \tag{2.3}$$

where the temporal wavefunction obeys,

$$i\frac{d\phi(t)}{dt} = E\phi(t), \tag{2.4}$$

and therefore oscillates as,

$$\phi(t) = \phi(0)e^{-iEt}, \tag{2.5}$$

and where the spatial wavefunction obeys the time-independent Schrödinger Equation,

$$\hat{H}\psi(\mathbf{x}) = E\psi(\mathbf{x}), \tag{2.6}$$

where E is the constant energy associated with the wavefunction. Broadly speaking, this thesis is concerned with the development of new computational methods to find approximate solutions to this equation for atoms and molecules.

Specifying the Hamiltonian, $\hat{H}$, fully defines the quantum-mechanical problem, that is, the eigenproblem that must be solved in order to obtain the wavefunction, and for the case of an arbitrary molecule consisting of N electrons and M nuclei, the Hamiltonian is,

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \frac{1}{2}\sum_{a=1}^{M}\frac{1}{m_a}\nabla_a^2 - \sum_{i=1}^{N}\sum_{a=1}^{M}\frac{Z_a}{r_{ia}} + \sum_{i=1}^{N}\sum_{j=1}^{i-1}\frac{1}{r_{ij}} + \sum_{a=1}^{M}\sum_{b=1}^{a-1}\frac{Z_aZ_b}{r_{ab}}. \tag{2.7}$$

This, of course, implies that not just the electrons, but also the nuclei should be treated with a wave description. It is more intuitive to force the separability of the electronic and nuclear wavefunctions, and to model the nuclei as classical particles that therefore have well-defined positions and momenta. The clamped nuclei picture can then be invoked in which the nuclei are held fixed and act only as parameters defining the potential for the electronic problem. Then, one need only solve the Schrodinger equation for the electrons, for which the Hamiltonian is,

$$\hat{H}_{\text{elec}} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\sum_{a=1}^{M}\frac{Z_a}{r_{ia}} + \sum_{i=1}^{N}\sum_{j=1}^{i-1}\frac{1}{r_{ij}}. \tag{2.8}$$

This approximation, of course, does not reflect the true quantum-mechanical situation, but often works exceptionally well in practice because the mass of an electron is roughly one thousand eight hundred times smaller than the mass of the simplest atomic nucleus (the single proton of a hydrogen atom). The electrons therefore "move" much faster than the nuclei, and, to an electron, the nuclear motion indeed appears to be fixed, or, stated another way, the electrons are able to respond almost instantaneously to changes in the nuclear positions. This is all part of the well-known Born-Oppenheimer approximation, which will be assumed throughout this thesis, the other part of which concerns the solution for the nuclear wavefunction in the field of the electrons. Corrections to the Born-Oppenheimer approximation have been devised, but it will not become necessary to apply them in this work.

The time-independent Schrödinger Equation is mathematically intractable to exact analytic solution, even in the regime of the Born-Oppenheimer approximation, except for the simplest of systems, such as atomic hydrogen. A huge number of schemes, sometimes called quantum-chemical methods or *ab initio* methods, have therefore been devised to find approximate numerical solutions to the Born-Oppenheimer electronic Schrödinger equation using computational tools, and the most popular will be highlighted in this chapter. The simplest of these, Hartree-Fock theory, is used by a great many *ab initio* methods, the so-called single-reference methods, as a first approximation that is then corrected upon. Hartree-Fock theory will be introduced shortly, motivated by a discussion

of the quantum mechanics of identical particles in the following section.

2.2 Identical Particles

In classical physics, it is very common to make the simplifying assumption that certain particles are exactly identical. For example, modelling the collision of balls on a snooker table becomes much easier if all the balls are assumed to be perfect spheres of the same uniform density and fixed radius. It is, of course, understood that the balls are not strictly identical in practice; small deviations may occur within each ball, and between balls, and it is possible to attach unique labels to each of the balls. When quantum particles are under consideration, however, this is not true at all. The electron, for example, is a fundamental and indivisible particle, at least to contemporary particle physics, and (ignoring spin) every electron in the universe is strictly identical with every other electron in the universe. There is no way to attach a unique label to an electron, even in principle. This indistinguishability of quantum particles has great ramifications for the behaviour of the wavefunction under particle interchange, as will be discussed in this section.

Consider a system consisting of two identical particles, and denote the wavefunction of this system as $\psi(1,2)$. The physically relevant quantity is the wavefunction square norm, $|\psi(1,2)|^2 = \bar{\psi}(1,2)\psi(1,2)$. Since the two particles are identical, exchanging one particle for the other cannot have an empirically observable effect, and therefore this quantity cannot be affected by this interchange,

$$|\psi(1,2)|^2 = |\psi(2,1)|^2. \tag{2.9}$$

This equation, which reduces to,

$$|\psi(1,2)| = |\psi(2,1)|, \tag{2.10}$$

has the solutions,

$$\psi(1,2) = \pm\psi(2,1). \tag{2.11}$$

If the wavefunction is symmetric under particle interchange, $\psi(1,2) = +\psi(2,1)$, the associated particles are called bosons. A photon is a well-known example of a boson. If the wavefunction is instead antisymmetric, $\psi(1,2) = -\psi(2,1)$, the particles are called fermions. The electron is a fermion, and it is therefore this type of particle that will be of interest in this thesis.

Consider next an example of non-interacting fermions, such as two electrons separated by a distance such that their electrostatic interactions become negligible. Since the electrons are isolated from one another, they behave as independent particles. This implies that the two-electron wavefunction, $\psi(1,2)$, must be separable into contributions from the individual electrons,

$$\psi(1,2) = \psi_a(1)\psi_b(2). \tag{2.12}$$

Since the wavefunction square norm is often, at least in the Born interpretation of Quantum Mechanics, thought of as a probability density function, basic probability theory explains why the form of the separation is a product, and not a sum ("1 *and* 2" not "1 *or* 2"). This wavefunction is not quite correct, however, because it does not yet satisfy the fermionic antisymmetry property. Similarly, since the electrons are identical, it could be "electron 2" in ψ_a, and "electron 1" in ψ_b. This suggests that the 2-electron wavefunction is better written as the antisymmetric linear combination of these two possibilities,

$$\psi(1,2) = \psi_a(1)\psi_b(2) - \psi_b(1)\psi_a(2). \tag{2.13}$$

Strictly, the opposite-sign wavefunction,

$$\psi(1,2) = \psi_b(1)\psi_a(2) - \psi_a(1)\psi_b(2) = -(\psi_a(1)\psi_b(2) - \psi_b(1)\psi_a(2)), \tag{2.14}$$

is also valid, but since they differ only by a sign change, they have identical square norms and are thus physically indistinguishable. Consider next what happens if the two electrons occupy the same state,

$$\psi_a = \psi_b. \tag{2.15}$$

In this case, the two-electron wavefunction vanishes,

$$\psi(1,2) = \psi_a(1)\psi_a(2) - \psi_a(1)\psi_a(2) = 0, \tag{2.16}$$

such that this occurs with zero probability, that is, it is impossible. This is the well-known Pauli Exclusion Principle.

Finally, of interest is how the above results generalize to 3, 4, or an arbitrary number of independent electrons. Perceptive readers may already have noted that the two-electron wavefunction given above may, in fact, be written as a 2×2 (Slater) determinant,

$$\psi(1,2) = \psi_a(1)\psi_b(2) - \psi_b(1)\psi_a(2) = \begin{vmatrix} \psi_a(1) & \psi_b(1) \\ \psi_a(2) & \psi_b(2) \end{vmatrix} \tag{2.17}$$

and this allows trivial generalization to a 3-electron wavefunction, as a 3×3 determinant,

$$\psi(1,2,3) = \begin{vmatrix} \psi_a(1) & \psi_b(1) & \psi_c(1) \\ \psi_a(2) & \psi_b(2) & \psi_c(2) \\ \psi_a(3) & \psi_b(3) & \psi_c(3) \end{vmatrix} \tag{2.18}$$

and to wavefunctions for any number of independent fermions; an N-electron independent fermion wavefunction can be written as an N×N determinant, for example. The fermionic antisymmetry property, that swapping any pair of electrons should change the sign of the wavefunction, is automatically captured through the mathematics of elementary determinantal row and column operations[60].

Although electrons are, in general, not independent of each other in real situations such as atoms, molecules or solids, due to their electrostatic interactions, an independent particle model can often be an excellent starting approximation for investigating the physics of such systems, such as in Hartree-Fock theory, as will be discussed shortly. Furthermore, Slater determinants play a pivotal role in more advanced electronic structure models that attempt to go beyond the independent particle approximation by the inclusion of electron correlation, such as in both Configuration Interaction and Coupled Cluster theories, and the associated fermionic antisymmetry property will additionally be inherited by the amplitudes (coefficients) used in the determinantal wavefunction expansion of each theory. For this reason, it is worthwhile to mention the important relationships satisfied by Slater determinants at this stage, although more sophisticated discussions and proofs can be found elsewhere[59]. In particular, two Slater determinants, $|\Phi_1\rangle$

and $|\Phi_2\rangle$, which differ in any of the 1-electron wavefunctions (otherwise known as occupied spinorbitals) from which they are constructed, $\psi_a, \psi_b, \ldots$, automatically have zero overlap,

$$\langle \Phi_1 | \Phi_2 \rangle = 0, \tag{2.19}$$

and when these Slater determinants are later viewed as basis vectors for the exact many-electron wavefunction, this ensures that linearly independent Slater determinants are automatically mutually orthogonal. Furthermore, certain matrix elements between different Slater determinants also vanish. If $\hat{O}$ is a 1-electron operator, and $|\Phi_1\rangle$ and $|\Phi_2\rangle$ differ by more than a single spinorbital, then the matrix element $\langle \Phi_1 | \hat{O} | \Phi_2 \rangle$ vanishes. Similarly, if $\hat{O}$ is a 2-electron operator, and $|\Phi_1\rangle$ and $|\Phi_2\rangle$ differ by more than two spinorbitals, then the matrix element $\langle \Phi_1 | \hat{O} | \Phi_2 \rangle$ also vanishes. Since the Coulomb interaction acts between pairs of charged particles, the molecular Hamiltonian, the subject of the following subsection, contains only 1-electron (for example, the kinetic energy operator) and 2-electron operators.

2.3 The Hartree-Fock Approximation and the Introduction of a Basis

The previous subsection established that the wavefunction for a system of independent fermions takes the form of a Slater determinant. The ansatz of *Hartree-Fock* (HF) theory[1, 2] is that a Slater determinant is an adequate model of interacting fermions also. This invokes a mean-field independent-particle approximation for the interacting fermions, as will be discussed below. In order to compute a ground-state electronic structure with Hartree-Fock theory, the fundamental principle is that the set of spin-orbitals, from which the single-determinantal wavefunction is constructed, should be adjusted according to the variational principle until the determinant of minimal energy is constructed. This determinant is the ground-state Hartree-Fock wavefunction, and its energy is the ground-state Hartree-Fock energy, an approximation of the true energy of the system.

The ground-state Hartree-Fock energy associated with the single-determinantal Hartree-Fock wavefunction $|\Phi_0\rangle$ is given by the standard quantum-mechanical energy expectation value expression,

$$\begin{aligned} E_{\mathrm{HF}} &= \frac{\langle\Phi_0|\hat{H}|\Phi_0\rangle}{\langle\Phi_0|\Phi_0\rangle} \\ &= \langle\Phi_0|\hat{H}|\Phi_0\rangle, \end{aligned} \tag{2.20}$$

where the convention that Slater determinants be normalized to unity has been applied. $|\Phi_0\rangle$ is not a variable, but a function of the electronic coordinates. The above energy depends only on the form taken by $|\Phi_0\rangle$, and is therefore not a function, but a *functional*. The mathematical mechanism for the minimization of such a functional is called *Calculus of Variations*, and is well described in mathematics textbooks[61, 62], and also in the context of Hartree-Fock theory[58, 59]. The full analysis is beyond the scope of this chapter, but the key points can be summarized as follows. Hartree-Fock theory constrains the wavefunction to take the form of a single Slater determinant, and this form is therefore not free to vary during the above minimization. Instead, it is the form of the 1-electron wavefunctions, the spinorbitals, here denoted as $\{\chi_i\}$, that is allowed to vary to minimize the above Hartree-Fock functional. The minimization is performed with respect to the constraint that the spinorbitals remain orthonormal, and a Lagrangian is therefore constructed. The necessary condition for a minimum, subject to the orthonormality constraint, is that the variation in the Lagrangian due to a small variation in the spinorbitals should vanish, exactly as the vanishing of the gradient of a function of one variable is the condition for a stationary point in differential calculus. This condition leads to the following set of integro-differential equations, the Hartree-Fock equations, that must be solved in order to determine the optimum form of the 1-electron wavefunctions that yield the minimal Hartree-Fock energy,

$$\hat{f}|\chi_i\rangle = \sum_j \epsilon_{ji}|\chi_j\rangle, \tag{2.21}$$

where ϵ is the matrix of Lagrange multipliers, and where $\hat{f}$ is the Fock operator,

$$\hat{f} = \hat{h}(1) + \sum_j \hat{\mathcal{J}}_j(1) - \sum_j \hat{\mathcal{K}}_j(1), \tag{2.22}$$

which consists of the core Hamiltonian,

$$\hat{h}(1) = -\frac{1}{2}\nabla_1^2 - \sum_a \frac{Z_a}{r_{1a}}, \tag{2.23}$$

the Coulomb operators,

$$\hat{\mathcal{J}}_i(1)\chi_j(1) = \chi_j(1)\int \bar{\chi}_i(2) r_{12}^{-1} \chi_i(2) d\mathbf{x}_2, \tag{2.24}$$

and the Exchange operators,

$$\hat{\mathcal{K}}_i(1)\chi_j(1) = \chi_i(1)\int \bar{\chi}_i(2) r_{12}^{-1} \chi_j(2) d\mathbf{x}_2. \tag{2.25}$$

However, a simplification can be made by noting that the Hartree-Fock energy is invariant to a unitary transformation of the spinorbitals amongst themselves; exploiting this allows the HF equations to be recast into their canonical form,

$$\hat{f}|\chi_i\rangle = \epsilon_i|\chi_i\rangle. \tag{2.26}$$

The problem of finding the single Slater determinant of minimal energy is therefore reduced to the problem of solving this set of equations.

At this point, it is necessary to point out that this discussion of the Hartree-Fock approximation has been motivated by the desire to find an approximate solution to the N-electron Born-Oppenheimer electronic Schrödinger equation, but it is also legitimate to ask for what Hamiltonian Hartree-Fock theory yields the correct eigenstate. That is, for what problem is Hartree-Fock the exact answer? The solved Hamiltonian is, in fact, the Hamiltonian that results by the replacement of the instantaneous inter-electronic interactions by an average acquired by integration over the electronic coordinates, and the Hartree-Fock approximation therefore corresponds to an independent-electron model in which each electron moves in the average (or mean) field of the other electrons. For a much more thorough discussion of Hartree-Fock theory, the reader is directed to, for example, references [58] or [59].

Of course, computers are best at performing numerous arithmetic operations extremely quickly, and the Hartree-Fock equations above are therefore not well-posed for computational implementation. Instead, the problem of solving these equations must be converted to a problem of linear algebra. In order to do

this, a *basis set*, is introduced. That, is, each orbital, an unknown function of the coordinates of a single electron prior to the solution of the Hartree-Fock equations, is expanded as a linear combination of a set of K *known* functions, the basis functions, with undetermined coefficients,

$$\chi_i = \sum_{\mu=1}^{K} C_{\mu i}\phi_\mu \quad i = 1, 2, \ldots K \tag{2.27}$$

This reduces the problem of finding the spinorbital functions to the problem of finding the set of discrete amplitude coefficients, a problem that is tractable to computational implementation. Manipulating the Hartree-Fock equations further and inserting this expansion yields the Roothaan equations,

$$\mathbf{FC} = \mathbf{SC}\epsilon \tag{2.28}$$

a matrix eigenproblem, to be solved for the matrix of coefficients, $\mathbf{C}$ and the matrix of eigenvalues ϵ, where,

$$F_{\mu\nu} = \langle\phi_\mu|\hat{f}|\phi_\nu\rangle \tag{2.29}$$

$$S_{\mu\nu} = \langle\phi_\mu|\phi_\nu\rangle. \tag{2.30}$$

Most basis sets consist of Gaussian functions, which facilitate simple integral evaluation. In most basis sets, the number of basis functions is significantly greater than the number of occupied orbitals, and the Hartree-Fock-Roothaan equations therefore also generate a set of unoccupied (or virtual) orbitals. These are extremely useful for correlated wavefunction methods, as will be discussed shortly.

Such a basis set expansion is exact only if the basis is complete, which would unfortunately require infinitely many basis functions ($K \to \infty$), which would itself remain intractable to computational application. Instead, a finite basis set must be employed, with the criterion that as the size of the basis set is made larger, it should tend towards completeness. It is noteworthy that the correlation energy, to be discussed in detail in the following section, predicted by more advanced quantum-chemical methods, does not converge quickly with the size of the basis. This was possibly first noted by Hylleraas, while investigating orbital expansions of two-electron systems[63]. Acquiring chemical accuracy from basic

calculations can therefore require very large basis sets, at great computational expense. For this reason, techniques for extrapolation of the correlation energy to the complete basis limit may be deployed, and there has also been much recent work on the development of explicitly correlated wavefunction methods that improve the rate of basis set convergence[64].

The calculations in this thesis use either the minimal STO-3G basis when comparison of results with more accurate (and more computationally demanding) methods is required, or make use of the correlation-consistent basis sets of Dunning and co-workers[65] possibly augmented with additional diffuse functions[66], for more reliable benchmarking.

2.4 Single- and Multi-Reference Post-Hartree-Fock Methods

Hartree-Fock theory normally works exceptionally well, and typically accounts for 99% of the total energy of a molecule. Unfortunately, as was discussed above, the Hartree-Fock single-determinantal wavefunction ansatz corresponds to a mean-field treatment of the electrons, and the small remaining uncaptured energy, termed the *correlation energy* since it is the energy associated with the correlated motion of the electrons, is itself extremely important for a correct description of, for example, chemical bonding. It is not precisely true to say that Hartree-Fock is a totally uncorrelated theory, however, because requiring the wavefunction take the form of a Slater determinant guarantees fermionic antisymmetry and therefore also that the Exclusion Principle holds. The exchange correlation energy included in the Hartree-Fock energy is a direct result of fermionic antisymmetry; in a purely uncorrelated Hartree N-electron wavefunction, a simple product of 1-electron wavefunction, the exchange term does not enter. Nevertheless, the Hartree-Fock approximation provides an inadequate description of chemical bonding, due to the *dynamic correlation* of the electrons. Many of the Post-Hartree-Fock methods to be discussed shortly are well-equipped to deal with dynamic correlation, however. The ubiquitous and extremely suc-

cessful CCSD(T) method, for example, typically describes molecular equilibria well within chemical accuracy.

However, a more serious breakdown of Hartree-Fock theory can occur if the approximation of a single-determinantal reference wavefunction ansatz becomes invalid and more than a single determinant becomes equally important to the description of the ground-state electronic structure. This is called *non-dynamic correlation*, or *static correlation*, and the single-reference methods, which each attempt to compute corrections to the Hartree-Fock approximation by some scheme, typically struggle much more when non-dynamic correlation becomes strong. This often happens in bond-breaking situations, especially of multiply-bonded molecules.

The problem is that the single-reference methods, taking truncated Configuration Interaction or Coupled Cluster theories (to be discussed shortly) as examples, explore only a subset of the full space of determinants. This subset is typically the set containing the reference determinant and those excited determinants that can be obtained from single and double excitations of the reference. In any system, some determinants are inevitably more important to a correct description of the electronic structure than others. If the most important determinants belong to the explored space, then the method can be expected to perform well, but otherwise can be expected to perform less well. In such situations, the most obvious solution is the construction of a reference wavefunction containing multiple determinants, thus potentially avoiding the non-dynamic breakdown of Hartree-Fock theory entirely, and then to employ a singles and doubles excitation scheme of those multiple included determinants in order to account for the effects of dynamic electron correlation. Unfortunately, the process of choosing the reference determinants relevant to the problem of interest is not always straightforward, especially for large systems, and often requires an intuitive grasp on the chemistry of the problem prior to running the calculation. This is a disadvantage of multireference methods, when contrasted with the simplicity of Hartree-Fock theory and methods based on it, for which the reference determinant is trivially defined as the determinant of minimal energy, and which can therefore be determined automatically, through the schemes described above. Many current multireference

methods also suffer from problems due to a lack of rigorous extensivity and unfavourable computational scaling, and these properties will be discussed in more detail below. In this thesis, multireference calculations may sometimes be deployed in lieu of more expensive FCI calculations, and used as the benchmark for a correct description of a system, in order to measure the effectiveness of other, less robust but more computationally practical methods, especially to gauge the relative performance of the single-reference methods in multireference situations. The multireference methods used for this task include internally-contracted multireference configuration interaction[29, 30] (MRCI) and multireference averaged quadratic coupled cluster[32] (MRAQCC).

The central topic of this thesis, however, is instead the possibility of constructing a new family of single-reference (and therefore "black-box") methods that remain robust to the breakdown of the Hartree-Fock approximation not just to include dynamic correlation of the electrons, but also to correctly describe problems for which non-dynamic correlation becomes strong. Before discussing this topic further, it is worthwhile to review the contemporary approaches to the inclusion of electron correlation, and this is best accomplished first by a discussion of the various methodological properties that each Post-Hartree-Fock method may or may not satisfy.

A Post-Hartree-Fock (PHF) method is, loosely, any *ab initio* quantum-chemical method that models, or attempts to approximately model, the effects of electron correlation omitted by Hartree-Fock theory. These methods therefore recover, or at least partially recover, the correlation energy, and are therefore, in principle, capable of modelling phenomena such as chemical bonding. A great variety of PHF methods have been suggested in the literature, from Møller-Plesset Perturbation theory and Configuration Interaction theory to Coupled Electron Pair methods and the extremely successful Traditional Coupled Cluster theory.

The major differences between these methods, which are each discussed in more detail in the following sections, is in the mathematical form that each assumes for the electronic wavefunction, the so-called electronic structure ansatz. Møller-Plesset theory, for example, assumes that the true electronic wavefunction, or at least the optimum wavefunction for the finite basis set in use, can be

expanded in orders of perturbation theory, with the Hartree-Fock determinant acting as the zeroth-order wavefunction, whereas Configuration Interaction and Coupled Cluster theories instead expand the exact wavefunction as a determinantal series directly. The theoretical distinctions between the different ansätze are best understood through the methodological properties implied by each ansatz and the subsequent ramifications on the level of accuracy with which each method can produce physically and chemically relevant results across a wide range of phenomena.

As such, this section summarizes some of the important methodological properties that can be possessed by the PHF methods. In principle, satisfying all of these properties should be a sufficient condition for an accurate and widely-applicable theory. In practice, however, a given PHF method will possess some, but not all of these properties. This list discusses the majority of the relevant properties, but is not exhaustive. For example, satisfaction of the generalized Hellmann-Feynman theorem[67], useful for the calculation of molecular properties, is only mentioned a few times in this thesis, and is not given further discussion here.

- **An exact treatment of limiting systems**

As was discussed above in the context of Hartree-Fock theory, real calculations must be performed in finite basis sets (sometimes called the finite 1-electron basis), since even the computational resources of the most powerful supercomputer cluster are never infinite. This means that it is never possible to obtain the exact ground-state solution to the Born-Oppenheimer electronic Schrödinger Equation, although large basis sets combined with basis set extrapolation techniques typically yield approximate solutions that approach and converge to the true solution well.

It *is* possible, however, to find the optimum approximate solution within any finite basis. This is because, as will be discussed shortly, this optimum wavefunction can be represented as a determinantal expansion (sometimes called the N-electron basis), which is itself finite in a finite 1-electron basis. This optimum

solution goes under various names, with Full Configuration Interaction (FCI) the most common, but Full Coupled Cluster (FCC) also used. Regardless of the name, however, the principle is that the finite N-electron solution space can, in principle, be fully explored, with FCI and FCC representing different parameterisations of the same complete space. The determinants are thus basis vectors for this space, and the optimum wavefunction is a superposition (a linear combination in the case of FCI) of each of these basis vectors.

Unfortunately, although it might be *possible* to perform an FCI calculation, it is so computationally prohibitive and impractical that it is very rarely even attempted, except for very small systems and with small 1-electron basis sets. Instead, it is typical to explore only a restricted subspace (or manifold) of the N-electron space. The question, then, is whether this can be a good approximation, but it gains credibility if it becomes exactly equivalent to FCI in some limiting cases. Of particular importance is the limit of 2 electrons; if a quantum-chemical method is exactly equivalent to FCI (or simply 'exact') for 2 electrons it guarantees, in a sense, that pairwise correlations are correctly treated, which can be assigned the highest priority given that the Coulomb force between electrons is itself a pairwise interaction.

- **Rigorous extensivity**

An extensive property is one which scales with system size. A simple hypothetical example from the field of thermodynamics is of a container of water; if the volume of the water is doubled then the total internal energy of the system doubles, because there is then twice as much matter. The internal energy is therefore a thermodynamically extensive quantity. In contrast, a thermodynamically intensive quantity is one that is invariant to the size of the system. In the previous example, the temperature of the system does not change upon the addition of more (identically prepared) water, and is therefore thermodynamically intensive.

In the context of quantum chemistry, the property of rigorous extensivity is taken to mean *correct* (that is linear) scaling of calculated energies with some measure of the size of the system. If a quantum-chemical method does not possess

this property, then, even if it is exceptionally accurate for some system, it does not necessarily achieve the same level of accuracy for the containing supersystem. Rigorous extensivity is thus an essential property that must be satisfied in order to treat extended systems reliably.

A similar, but perhaps more transparent property is that of size consistency. A quantum-chemical method is said to be size consistent if the calculated energy of a system containing two isolated (infinitely separated) subsystems is equal to the sum of the energies of those subsystems. This is an especially important property of the chosen *ab initio* method if the chemical system of study involves any form of dissociation or bond breaking, but is also useful as an indicator of extensivity; a method cannot be extensive unless it is size consistent. Extensivity, however, is a stronger property than size consistency, since rigorous extensivity implies size consistency, but the converse is not true.

There is an extremely useful result that, put simply, states that if the energy calculated by an approximate quantum mechanical method can be written in terms of fully linked diagrams only, diagrams that are themselves composed of only a single closed diagram, then this is equivalent to rigorous extensivity. The presence of even a single uncancelled unlinked diagram, however, destroys the property of rigorous extensivity, and, at best, the energy may be only approximately extensive, or, at worst, not extensive at all. The use of diagrammatic notation, particularly in the context of Coupled Cluster theory, is discussed at the end of this chapter.

- **Invariance to orbital rotations**

The word 'scalar' is commonly used to mean a quantity with magnitude, but no direction. To the mathematically-inclined, however, a scalar is much more precisely defined as a quantity that is invariant to rotations of the coordinate system in use[61, 62]. It has already been discussed how the solution to the Hartree-Fock-Roothaan equations is not unique, with the canonical Hartree-Fock orbitals chosen as those that diagonalize the Fock matrix, and that a rotation of the occupied or virtual spin-orbitals amongst themselves does not change the energy.

The Hartree-Fock energy is hence a scalar with respect to rotations in the orbital space $\{\psi_i\}$, and (since the Hartree-Fock energy does not depend upon it) also in the space $\{\psi_a\}$. It is, of course, not invariant to rotations between the occupied and virtual spaces, since such a rotation would generate an excited determinant, which, by definition, would be higher in energy than the Hartree-Fock reference determinant. This same invariance to rotations in the underlying orbital spaces $\{\psi_i\}$ and $\{\psi_a\}$ is demanded of correlated methods also, which would otherwise be dependent on the choice of Hartree-Fock orbitals. In fact, this invariance of the correlation energy to rotations in these orbital spaces allows a localized orbital basis to be chosen without affecting the calculated energy, facilitating a more local treatment of electron correlation, as in the local correlation methods such as LCCSD[68].

- **Definition through minimization of a functional**

Taylor's Theorem is that an infinitely-differentiable function may be represented as a series through the values of those infinitely-many derivatives at a single point. For the case of a function of a single variable, this may be concisely written as,

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n = f(a) + f'(a).(x-a) + \frac{f''(a)}{2!}.(x-a)^2 + \ldots, \tag{2.31}$$

and may be extended further to multivariate functions. This theorem may be applied to an analysis of the error present in the calculated energy of a quantum-chemical method, and if the energy of such a quantum-chemical method is defined to be the minimum of a functional, then a necessary condition for that minimum is that the first partial derivatives of the energy with respect to the constituent parameters vanish. This means that, at the minimum, the term containing the first derivative vanishes, and the error is at least second-order in the remaining errors in the parameters. These parameters are often called "cluster amplitudes" in the methods discussed later, and this property is a great advantage of functional-based methods over alternative approaches, such as those based on some projective scheme, for which errors are first-order in the parameter errors.

- **A variational upper bound on the exact ground-state Schrödinger eigenvalue**

Related to the concept of defining an *ab initio* method through the variational minimization of a functional is the property of a variational upper bound. Consider calculating the ground-state energy associated with some approximate wavefunction $|\Phi\rangle$ through direct insertion into and minimization of the quantum mechanical expectation value expression,

$$E = \frac{\langle\Phi|\hat{H}|\Phi\rangle}{\langle\Phi|\Phi\rangle}. \tag{2.32}$$

Since this is only an approximate wavefunction, this energy need not equal the exact ground-state Schrödinger eigenvalue. To gain additional insight, let $|k\rangle$ denote the (unknown) eigenstates of the Hamiltonian, for $k = 0, 1, 2, \ldots$ and let E_k denote the associated energies, such that the Schrödinger equation for each eigenstate is,

$$\hat{H}|k\rangle = E_k|k\rangle. \tag{2.33}$$

Inserting the resolution of the identity,

$$E = \frac{\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\langle\Phi|k\rangle\langle k|\hat{H}|l\rangle\langle l|\Phi\rangle}{\sum_{k=0}^{\infty}\langle\Phi|k\rangle\langle k|\Phi\rangle}, \tag{2.34}$$

then applying the Schrödinger equation and noting that the eigenstates of the Hamiltonian can be taken to be orthonormal,

$$E = \frac{\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\langle\Phi|k\rangle E_l\langle k|l\rangle\langle l|\Phi\rangle}{\sum_{k=0}^{\infty}\langle\Phi|k\rangle\langle k|\Phi\rangle} = \frac{\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\langle\Phi|k\rangle E_l\delta_{kl}\langle l|\Phi\rangle}{\sum_{k=0}^{\infty}\langle\Phi|k\rangle\langle k|\Phi\rangle} = \frac{\sum_{k=0}^{\infty}E_k\langle\Phi|k\rangle\langle k|\Phi\rangle}{\sum_{k=0}^{\infty}\langle\Phi|k\rangle\langle k|\Phi\rangle}. \tag{2.35}$$

Finally, noting that, by definition, the ground-state is the lowest energy state, followed by the first excited state and so on,

$$E_0 \geq E_1 \geq E_2 \geq \ldots, \tag{2.36}$$

it follows that,

$$E = \frac{\sum_{k=0}^{\infty} E_k \langle \Phi|k\rangle\langle k|\Phi\rangle}{\sum_{k=0}^{\infty} \langle \Phi|k\rangle\langle k|\Phi\rangle} \geq \frac{\sum_{k=0}^{\infty} E_0 \langle \Phi|k\rangle\langle k|\Phi\rangle}{\sum_{k=0}^{\infty} \langle \Phi|k\rangle\langle k|\Phi\rangle} = E_0 \frac{\sum_{k=0}^{\infty} \langle \Phi|k\rangle\langle k|\Phi\rangle}{\sum_{k=0}^{\infty} \langle \Phi|k\rangle\langle k|\Phi\rangle} = E_0, \quad (2.37)$$

and the energy associated with the approximate wavefunction is hence always greater than or equal to E_0; it is an upper bound on the exact ground-state Schrödinger energy eigenvalue, or, stated another way, this eigenvalue is a lower bound on the energy calculable by the approximate method.

In practice, if an electronic structure method possesses this variational upper bound property, it means that its calculated potential energy curves will remain above and not cross the FCI curve. A central theme in this thesis is that the highly successful Traditional Coupled Cluster method abandons this property in favour of rigorous extensivity, but that this causes non-variational divergent behaviour to occur when non-dynamic correlation becomes strong, and that a method that preserves all of the previously mentioned methodological properties while also approximately satisfying the variational upper bound property is instead what is required.

- **Satisfactory computational requirements**

A final criterion for a good quantum-chemical method is that it be computationally practical; if computational feasibility were not an issue, there would be no reason to even consider alternative schemes to FCI. In practice, the computational demands of any algorithm may be specified by two measures. The first of these measures is the computaional complexity, also known as time complexity, or, more loosely, just the computational cost. This measures the total number of numerical operations that a computer must carry out in order to complete the algorithm, or, more simply, how long one can expect the completion of the algorithm to take. The second measure is the memory, or storage requirement, of the algorithm, that is, how much data must be stored in memory, and the algorithm have access to, at any one time. Both measures are typically classified by their limiting step, using a Big-O, $\mathcal{O}$, notation. For example, in this thesis, an

ab initio method will be taken to be computationally practical if its complexity is at most $\mathcal{O}(o^2v^4)$ for the iterative step of the calculation and at most $\mathcal{O}(o^3v^4)$ for any non-iterative steps, and its storage requirements are at most $\mathcal{O}(o^2v^2)$. These are the computational requirements of the extremely popular and successful CCSD(T) method, to be discussed shortly. The o and v notation refers to the number of occupied and virtual orbitals in the calculation respectively, where, typically, $v >> o$.

There are many examples in the literature, such as the Quadratic Coupled Cluster method[55], which provide significantly more robust and accurate descriptions of ground-state molecular electronic strucures than CCSD, but which have not found widespread acceptance or application due to higher computational demands. In the case of Quadratic Coupled Cluster, its complexity, at $\mathcal{O}(v^6)$, is problematic. It is therefore essential that any new single-reference electronic structure ansätze that attempt to treat the electronic structure at a higher level than CCSD(T) do not violate this criterion for computational practicality by exceeding the specifications outlined above.

2.5 Møller-Plesset Perturbation Theory

The most obvious approach to the treatment of the electron correlation problem is to treat Hartree-Fock theory as a solved zeroth-order problem, and to compute corrections to Hartree-Fock through perturbation theory. In summary, the Møller-Plesset treatment of electron correlation partitions the Hamiltonian into two parts, $\hat{F}$, the Hartree-Fock Hamiltonian, the Hamiltonian that Hartree-Fock theory treats exactly (the sum of the Fock operators), and $\hat{V}$, the *Fluctuation Potential* that contains the difference between the instantaneous and mean-field Coulomb interactions, such that the true electronic Hamiltonian may be written as,

$$\hat{H} = \hat{F} + \hat{V}. \tag{2.38}$$

The energy and wavefunction may then be expanded in orders of perturbation theory as,

$$E = E^{(0)} + E^{(1)} + E^{(2)} + \dots \tag{2.39}$$

$$|\Phi\rangle = |\Phi^{(0)}\rangle + |\Phi^{(1)}\rangle + |\Phi^{(2)}\rangle + \dots, \tag{2.40}$$

Inserting these expansions into the Schrödinger equation and grouping by perturbation order, treating the fluctuation potential as first-order, the zeroth-order equation is,

$$\hat{F}|\Phi^{(0)}\rangle = E^{(0)}|\Phi^{(0)}\rangle \tag{2.41}$$

the first-order equation is,

$$\hat{F}|\Phi^{(1)}\rangle + \hat{V}|\Phi^{(0)}\rangle = E^{(0)}|\Phi^{(1)}\rangle + E^{(1)}|\Phi^{(0)}\rangle \tag{2.42}$$

the second-order equation is,

$$\hat{F}|\Phi^{(2)}\rangle + \hat{V}|\Phi^{(1)}\rangle = E^{(0)}|\Phi^{(2)}\rangle + E^{(1)}|\Phi^{(1)}\rangle + E^{(2)}|\Phi^{(0)}\rangle \tag{2.43}$$

and so on. Assuming a Hartree-Fock reference, from these, expressions for the zeroth-order energy,

$$E^{(0)} = \langle\Phi^{(0)}|\hat{F}|\Phi^{(0)}\rangle, \tag{2.44}$$

the first-order energy,

$$E^{(1)} = \langle\Phi^{(0)}|\hat{V}|\Phi^{(0)}\rangle, \tag{2.45}$$

and the second-order energy,

$$E^{(2)} = \langle\Phi^{(0)}|\hat{V}|\Phi^{(1)}\rangle, \tag{2.46}$$

may be obtained by projection onto the reference determinant. Given $|\Phi^{(0)}\rangle = |\Phi_0\rangle$, the sum of the first and second-order Møller-Plesset energies is the Hartree-Fock energy, and the first contribution to the correlation energy therefore enters at second order. The second-order energy depends on the first-order wavefunction, which can be determined by expanding it in terms of the exact eigenstates of the zeroth-order Hamiltonian, and then by inserting this expansion into the first-order equation, followed by projection onto those eigenstates to determine the

coefficients of the expansion. It is therefore possible to show that $E^{(2)}$ takes the following form,

$$E^{(2)} = \frac{1}{4}\sum_{ijab}\frac{\langle ij||ab\rangle\langle ab||ij\rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}. \tag{2.47}$$

where ϵ_i is defined through Eq. 2.26, and,

$$\langle ij||ab\rangle = \langle ij|ab\rangle - \langle ij|ba\rangle \tag{2.48}$$

$$\langle ij|ab\rangle = \int \bar{\chi}_i(\mathbf{x}_1)\bar{\chi}_j(\mathbf{x}_2)r_{12}^{-1}\chi_a(\mathbf{x}_1)\chi_b(\mathbf{x}_2)d\mathbf{x}_1 d\mathbf{x}_2. \tag{2.49}$$

Higher-order energies may be found similarly, and the sum of $E^{(0)}$, $E^{(1)}$, $E^{(2)}$ to some $E^{(N)}$ defines the MPN energy. For more information, the reader is directed to one of the numerous textbooks on the subject, such as reference [59].

While Møller-Plesset theory may seem like a sensible scheme for the inclusion of electron correlation effects, and while it is indeed widely used, especially at the MP2 level as a lowest-order correction to Hartree-Fock, it possesses two quite serious disadvantages. First, Møller-Plesset theory can only compute the exact electronic wavefunction for any system if the perturbation series is summed to infinite order, meaning that even simple two-electron systems cannot be described *exactly*. Second, the perturbation series may not converge at all[4–6], and instead diverge, calling into question the validity of the MP2, MP3, MP4, ... hierarchy as a series of systematically increasing accuracy. This precludes Møller-Plesset theory as the single-reference method of choice for robust and quantitatively accurate descriptions of molecular phenomena, and other schemes are necessary, although these more advanced methods often additionally employ perturbation theory in some capacity, such as for the inclusion of the effects of triple excitations, to be discussed later, in order to avoid undesirably high computational costs.

2.6 Configuration Interaction Theory

Instead of attempting to include the effects of electron correlation through perturbation theory, it is possible to show that the true N-electron wavefunction may be expanded exactly in the (finite) basis of all possible unique N×N Slater determinants formed from a complete set of spin orbitals (or basis functions) $\{\psi_i\}$.

For example, a 1-electron wavefunction may be expanded trivially in a basis as follows,

$$\Phi(\mathbf{x}_1) = \sum_i C_i \psi_i(\mathbf{x}_1), \tag{2.50}$$

which is automatically a determinantal expansion (of 1×1 determinants). A 2-electron wavefunction may be similarly expanded,

$$\Phi(\mathbf{x}_1, \mathbf{x}_2) = \sum_i \sum_j C_{ij} \psi_i(\mathbf{x}_1) \psi_j(\mathbf{x}_2), \tag{2.51}$$

then, requiring fermionic antisymmetry, $\Phi(\mathbf{x}_1, \mathbf{x}_2) = -\Phi(\mathbf{x}_2, \mathbf{x}_1)$,

$$\begin{aligned}
\sum_i \sum_j C_{ij} \psi_i(\mathbf{x}_1) \psi_j(\mathbf{x}_2) &= -\sum_i \sum_j C_{ij} \psi_i(\mathbf{x}_2) \psi_j(\mathbf{x}_1) \\
&= -\sum_i \sum_j C_{ji} \psi_j(\mathbf{x}_2) \psi_i(\mathbf{x}_1) \\
&= -\sum_i \sum_j C_{ji} \psi_i(\mathbf{x}_1) \psi_j(\mathbf{x}_2),
\end{aligned} \tag{2.52}$$

from which it follows that, for all i, j, $C_{ij} = -C_{ji}$, which, in particular, means $C_{ii} = 0$. Therefore, the 2-electron wavefunction may be written more compactly as,

$$\begin{aligned}
\Phi(\mathbf{x}_1, \mathbf{x}_2) &= \sum_i \sum_j C_{ij} \psi_i(\mathbf{x}_1) \psi_j(\mathbf{x}_2) \\
&= \sum_{i>j} C_{ij} \psi_i(\mathbf{x}_1) \psi_j(\mathbf{x}_2) + \sum_{i<j} C_{ij} \psi_i(\mathbf{x}_1) \psi_j(\mathbf{x}_2) \\
&= \sum_{i>j} C_{ij} \psi_i(\mathbf{x}_1) \psi_j(\mathbf{x}_2) - \sum_{i<j} C_{ji} \psi_i(\mathbf{x}_1) \psi_j(\mathbf{x}_2) \\
&= \sum_{i>j} C_{ij} \psi_i(\mathbf{x}_1) \psi_j(\mathbf{x}_2) - \sum_{j<i} C_{ij} \psi_j(\mathbf{x}_1) \psi_i(\mathbf{x}_2) \\
&= \sum_{i>j} C_{ij} \psi_i(\mathbf{x}_1) \psi_j(\mathbf{x}_2) - \sum_{i>j} C_{ij} \psi_i(\mathbf{x}_2) \psi_j(\mathbf{x}_1) \\
&= \sum_{i>j} C_{ij} \left(\psi_i(\mathbf{x}_1) \psi_j(\mathbf{x}_2) - \psi_i(\mathbf{x}_2) \psi_j(\mathbf{x}_1) \right) \\
&= \sum_{i>j} C_{ij} \begin{vmatrix} \psi_i(\mathbf{x}_1) & \psi_j(\mathbf{x}_1) \\ \psi_i(\mathbf{x}_2) & \psi_j(\mathbf{x}_2) \end{vmatrix}
\end{aligned} \tag{2.53}$$

This is therefore an expansion of the 2-electron wavefunction in the basis of all unique 2×2 Slater determinants that can be formed from the 1-electron orbital

basis. This discussion is given more treatment in various textbooks on electronic structure theory, such as Szabo and Ostlund[59].

The N-electron case can be proved by mathematical induction. In contrast to Many-Body or Møller-Plesset perturbation theory, which must go to infinite-order to account exactly for all of the correlation energy in a given 1-electron basis (assuming the perturbation series even converges), and which is hence computationally intractable, if the set $\{\psi_i\}$ is finite (corresponding to a finite 1-electron basis set), there are only finitely many determinants in the above expansion. Upon determination of the coefficients, this wavefunction therefore represents the optimal wavefunction for the finite basis. If the basis is allowed to approach completeness, this wavefunction can approach an exact eigenstate of the Born-Oppenheimer electronic Schrödinger Equation.

Furthermore, each of these finitely-many determinants may be related to the Hartree-Fock reference determinant; a Hartree-Fock scheme yields a set of M spin-orbitals, and these may be conceptually partitioned into the N occupied orbitals (those orbitals that are used to construct the Hartree-Fock single-determinantal ground-state wavefunction), $\{\psi_i, \psi_j, \psi_k, \ldots\}$, and the remaining M-N unoccupied (or virtual) spin-orbitals, $\{\psi_a, \psi_b, \psi_c, \ldots\}$. The set of all possible N×N Slater determinants for the finite 1-electron basis can therefore be generated by replacement of one or more of the occupied spin-orbitals by virtual orbitals. For example, if the reference determinant for a 2-electron system is denoted as,

$$|\psi_i\psi_j\rangle = \begin{vmatrix} \psi_i(1) & \psi_j(1) \\ \psi_i(2) & \psi_j(2) \end{vmatrix}, \tag{2.54}$$

then another possible determinant, also often termed an 'excited determinant' because the energy of this single-determinantal wavefunction must be higher than the optimal single-determinantal ground-state wavefunction by defintion, is,

$$|\psi_i\psi_a\rangle = \begin{vmatrix} \psi_i(1) & \psi_a(1) \\ \psi_i(2) & \psi_a(2) \end{vmatrix}. \tag{2.55}$$

If the Hartree-Fock reference determinant is $|\Phi_0\rangle$, then, for example, the notation $|\Phi_i^a\rangle$ refers to the excited determinant acquired by the replacement of orbital i with orbital a. Similarly, $|\Phi_{ij}^{ab}\rangle$ is the excited determinant acquired by replacement

of orbitals i and j with unoccupied orbitals a and b. The unoccupied orbitals come from the solution of the Hartree-Fock-Roothaan equations. With this notation, the linear determinantal expansion of Configuration Interaction theory can be more concisely written in terms of an excitation operator, $\hat{C}$, as,

$$|\Phi_{\mathrm{CI}}\rangle = (1 + \hat{C})|\Phi_0\rangle. \tag{2.56}$$

The $\hat{C}$ operator can itself be partitioned into contributions grouped by excitation rank,

$$\hat{C} = \hat{C}_1 + \hat{C}_2 + \ldots + \hat{C}_N, \tag{2.57}$$

where, using the Einstein Summation Convention[69], in which repeated indices imply summation,

$$\hat{C}_1|\Phi_0\rangle = C_a^i|\Phi_i^a\rangle, \tag{2.58}$$

and,

$$\hat{C}_2|\Phi_0\rangle = \frac{1}{4}C_{ab}^{ij}|\Phi_{ij}^{ab}\rangle, \tag{2.59}$$

and so on, where $\{C_a^i\} \cup \{C_{ab}^{ij}\} \cup \ldots$ is the set of amplitudes, or the coefficients of the determinants in the expansion.

In order to perform calculations through Configuration Interaction theory, the determinantal expansion can be inserted into the Schrödinger equation,

$$\hat{H}(1 + \hat{C})|\Phi_0\rangle = E_{\mathrm{CI}}(1 + \hat{C})|\Phi_0\rangle, \tag{2.60}$$

which can then be projected onto the reference determinant in order to determine an expression for the CI energy,

$$E_{\mathrm{CI}} = \langle\Phi_0|\hat{H}(1 + \hat{C})|\Phi_0\rangle, \tag{2.61}$$

and onto the manifolds of excited determinants in order to determine the equations to be solved for C_a^i,

$$\langle\Phi_a^i|\hat{H}(1 + \hat{C})|\Phi_0\rangle = E_{\mathrm{CI}}C_a^i, \tag{2.62}$$

for C_{ab}^{ij},

$$\langle\Phi_{ab}^{ij}|\hat{H}(1 + \hat{C})|\Phi_0\rangle = E_{\mathrm{CI}}C_{ab}^{ij}, \tag{2.63}$$

and so on. It is also possible to equivalently formulate Configuration Interaction theory through the minimization of a functional by insertion of the determinantal expansion into the quantum mechanical energy expectation value expression,

$$E_{\mathrm{CI}} = \frac{\langle \Phi_0 | (1 + \hat{C}^\dagger) \hat{H} (1 + \hat{C}) | \Phi_0 \rangle}{\langle \Phi_0 | (1 + \hat{C}^\dagger)(1 + \hat{C}) | \Phi_0 \rangle}, \tag{2.64}$$

and requiring that, invoking the variational principle, the optimum ground-state energy should be the minimum of this functional with respect to the CI amplitudes. This is called Variational Configuration Interaction, and will be discussed further in the following chapter.

If the excitation operator $\hat{C}$ allows for all possible excitations, that is, if $\hat{C}$ contains $\hat{C}_1$, $\hat{C}_2$, and so on, up to the number of electrons, the resulting theory is called Full Configuration Interaction (FCI), and, as has been discussed above, is the optimum approximate solution of the Born-Oppenheimer electronic Schrödinger equation within the finite basis set employed, and which converges to the exact solution if the basis approaches completeness. Unfortunately, the number of determinants in the expansion grows combinatorically with the number of basis functions, and therefore FCI is factorially complex, with the storage of the enormous number of amplitudes (one for each determinant) also a limiting factor. For these reasons, FCI calculations are typically utterly impractical for calculations on real systems of chemical interest, and are typically restricted to small molecules and small basis sets. However, the CI scheme presents an obvious avenue to the simplification of the calculation, which is the truncation of the $\hat{C}$ operator, and the subsequent restriction of the determinantal expansion, to include, for example, only single and double excitations of the reference determinant,

$$\hat{C} = \hat{C}_1 + \hat{C}_2. \tag{2.65}$$

This defines the CISD method, for which the computational complexity scales only as the sixth power of system size, allowing calculations to be performed on a significantly wider range of chemically relevant systems.

Unfortunately, truncated CI does not possess the property of size consistency, and, as a consequnce, cannot be extensive. This can be demonstrated quite

simply by consideration of the CI wavefunction for two isolated, non-interacting or otherwise independent subsystems, A and B,

$$\begin{aligned} \hat{C} &= \hat{C}_A + \hat{C}_B \\ |\Phi_0\rangle &= |\Phi_{0A}\rangle|\Phi_{0B}\rangle \\ (1+\hat{C})|\Phi_0\rangle &= (1+\hat{C}_A+\hat{C}_B)|\Phi_{0A}\rangle|\Phi_{0B}\rangle \\ &\neq (1+\hat{C}_A)|\Phi_{0A}\rangle.(1+\hat{C}_B)|\Phi_{0B}\rangle; \end{aligned} \tag{2.66}$$

the CI wavefunction is not separable, and does not correspond to the product of the CI wavefunctions of the monomers. Another perspective is that, in the equation for the doubles amplitudes coefficients in FCI and considering only double and quadruple excitations for simplicity,

$$\langle \Phi_{ab}^{ij}|\hat{H}(1+\hat{C}_2+\hat{C}_4)|\Phi_0\rangle = E_{\mathrm{CI}}C_{ab}^{ij}, \tag{2.67}$$

the operator $\hat{C}_4$ should be present on the left-hand-side. The resulting term approximately cancels with the right-hand-side, with the effect of eliminating unlinked terms that would otherwise enter the energy expression. Unfortunately, since quadruple excitations are totally omitted by CISD, this does not occur, and results in the loss of extensivity. Various schemes have been proposed that correct, or approximately correct, for the lack of extensivity in CI, such as the Quadratic CI[70] (QCI) method that takes a step towards Coupled Cluster theory (although QCI is rarely used today due to example calculations that have demonstrated the method to be less robust than Coupled Cluster[71]), the *a posteriori* corrections such as that of Davidson[17], and the Coupled Pair methods to be discussed in the following section.

2.7 Electron Pair Methods

As has just been explained, the CISD equations do not properly account for the effects of higher excitations, which leads to the presence of unlinked terms and the lack of rigorous extensivity. In the case of schemes such as Quadratic CI, which may be considered as an approximation to Coupled Cluster theory, to be discussed shortly, attempts are made to directly capture the effect of higher

excitations; quadruples for example, may be approximated as products of doubles, which resolves the problem. In the Coupled Electron Pair Approximation[10–15] (CEPA) methods, however, instead of attempting to build in the effects of higher excitations, simple modifications of the existing CISD equations are instead proposed. In the equation for the doubles cluster amplitudes in CISD,

$$\langle \Phi_{ab}^{ij} | \hat{H}(1+\hat{C}) | \Phi_0 \rangle = E_{\mathrm{CI}} C_{ab}^{ij}, \tag{2.68}$$

replacement of the energy E_{CI} by the Hartree-Fock reference energy, for example, yields the equation for CEPA(0). While simpler than CISD, CEPA(0) is not exact, even when the $\hat{C}$ operator is complete, for example (when $\hat{C} = \hat{C}_1 + \hat{C}_2$) for the case of a two-electron system. More complicated substitutions may be made to yield the CEPA(1), CEPA(2) and CEPA(3) methods, for example, and although these *are* exact for two electrons, unlike CISD and CEPA(0) (and also the Coupled Cluster methods), they are not strictly invariant to orbital rotations. For further information, there is an excellent discussion of the CEPA methods given in Reference [15].

Related to the CEPA methods are those methods that attempt to modify the Variational CI functional,

$$E_{\mathrm{CI}} = \frac{\langle \Phi_0 | (1+\hat{C}^\dagger) \hat{H} (1+\hat{C}) | \Phi_0 \rangle}{\langle \Phi_0 | (1+\hat{C}^\dagger)(1+\hat{C}) | \Phi_0 \rangle}, \tag{2.69}$$

which will itself be discussed in the following chapter where it is more relevant for introducing the overarching themes of this work. One well-known modification of the CI functional is the Coupled Pair Functional[8], which replaces the CI denominator, the part of the CI functional that leads to its lack of extensivity when $\hat{C}$ is truncated, with partial local denominators that have the effect of dividing the contributions to the CI numerator only by those parts of the denominator to which they are local. In the Averaged Coupled Pair Functional[31], the doubles contributions to the CI norm are instead simply divided by the number of electron pairs, rendering the denominator approximately intensive. Kollmar and Neese have also succeeded in crafting an extensive and orbitally-invariant energy functional that is approximately exact for 2-electron systems[72, 73], and recent work has resulted in some sophisticated new CEPA methods[18], and modifications of the Coupled Cluster method that may be understood through CEPA[74].

Furthermore, Mazziotti has done much work on *ab initio* quantum chemistry without wavefunctions[75, 76] that may be likened to CEPA, with impressive results[77, 78]. While CEPA methods therefore continue to show promise as electronic structure methods, they will be given little further consideration in this thesis, which will instead be focused on the extremely successful Coupled Cluster method[19–26].

2.8 Coupled Cluster Theory

From Configuration Interaction theory, it is known that the optimal wavefunction can be expanded in the basis of the Slater determinants that can be generated by the replacement of one or more of the orbitals occupied in the reference determinant with unoccupied (or virtual) orbitals,

$$|\Phi_{\mathrm{CI}}\rangle = (1 + \hat{C})|\Phi_0\rangle. \tag{2.70}$$

However, any truncation of this wavefunction, such as to include only single and double excitations of the reference determinant, leads to the lack of extensivity of the energy due to the presence of unphysical unlinked cluster contributions. This problem is, in fact, a direct result of the linear parameterisation of the CI wavefunction. Coupled Cluster[19, 20] theory proposes, instead, an exponential parameterisation,

$$|\Phi_{\mathrm{CC}}\rangle = e^{\hat{T}}|\Phi_0\rangle = \sum_{n=0}^{\infty} \frac{\hat{T}^n}{n!}|\Phi_0\rangle, \tag{2.71}$$

where meaning is given to this parameterisation through the Maclaurin series of the exponential function, as given above, and where it is convention to use the symbol $\hat{T}$ in place of $\hat{C}$.

Unlike the CI wavefunction, the CC wavefunction is separable for the case of

two isolated, non-interacting or otherwise independent subsystems,

$$\begin{aligned}
\hat{T} &= \hat{T}_A + \hat{T}_B \\
|\Phi_0\rangle &= |\Phi_{0A}\rangle|\Phi_{0B}\rangle \\
e^{\hat{T}}|\Phi_0\rangle &= e^{\hat{T}_A+\hat{T}_B}|\Phi_{0A}\rangle|\Phi_{0B}\rangle \\
&= e^{\hat{T}_A}e^{\hat{T}_B}|\Phi_{0A}\rangle|\Phi_{0B}\rangle \\
&= e^{\hat{T}_A}|\Phi_{0A}\rangle.e^{\hat{T}_B}|\Phi_{0B}\rangle,
\end{aligned} \tag{2.72}$$

which leads to size consistency, even for a truncated $\hat{T}$. This can be understood at a deeper level by realising that a truncated CC wavefunction approximately includes the effects of excitations higher than those included explicitly in the cluster operator. This is due to the non-linear nature of the parameterisation. For example, for the case of $\hat{T} = \hat{T}_1 + \hat{T}_2$, quadruple excitations enter through terms such as $\frac{1}{4!}\hat{T}_1^4$ and $\frac{1}{2!}\hat{T}_2^2$. Although this does not allow the weightings of the quadruples to vary independently of the singles and doubles, and there is therefore still an advantage in the explicit inclusion of triples and higher excitations into the cluster operator, such non-linear contributions exactly cancel the unphysical unlinked cluster contributions present in, for example, CISD. CC therefore achieves the same result as the QCI method, and the QCI method may, in fact, be considered an approximation to the CC method (or the CC method a generalization of QCI). In the limit of a complete cluster operator, the Full CI wavefunction and the Full CC wavefunction must agree, and the following relationships then exists between the single excitations,

$$\hat{C}_1 = \hat{T}_1, \tag{2.73}$$

the double excitations,

$$\hat{C}_2 = \hat{T}_2 + \frac{1}{2!}\hat{T}_1^2, \tag{2.74}$$

and similarly for all higher excitations.

In order to perform a calculation of, for example, a molecular energy within the CC scheme, the CC ansatz can be inserted into the Schrödinger Equation,

$$\hat{H}e^{\hat{T}}|\Phi_0\rangle = Ee^{\hat{T}}|\Phi_0\rangle, \tag{2.75}$$

which is then premultiplied by $e^{-\hat{T}}$,

$$e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle = E|\Phi_0\rangle, \tag{2.76}$$

which has the effect of decoupling the amplitude equations from the energy. This equation can then be projected onto the appropriate determinantal manifold in order to determine the working equations. In the case of a cluster operator restricted to only single and double excitations of the reference determinant, $\hat{T} = \hat{T}_1 + \hat{T}_2$, these equations, which define the CCSD method, are,

$$\langle\Phi_0|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle = E_{\text{CCSD}} \tag{2.77}$$

$$\langle\Phi_i^a|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle = 0 \tag{2.78}$$

$$\langle\Phi_{ij}^{ab}|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle = 0. \tag{2.79}$$

This projective approach is called Traditional Coupled Cluster[21–26] theory.

Since the Hamiltonian is at most a two-body operator, the Campbell-Baker-Hausdorff expansion[27] of the similarity-transformed Hamiltonian,

$$\bar{H} = e^{-\hat{T}}\hat{H}e^{\hat{T}} = \hat{H} + \left[\hat{H},\hat{T}\right] + \frac{1}{2!}\left[\left[\hat{H},\hat{T}\right],\hat{T}\right] + \frac{1}{3!}\left[\left[\left[\hat{H},\hat{T}\right],\hat{T}\right],\hat{T}\right] + \ldots \tag{2.80}$$

terminates at the fourth power of the cluster amplitudes[79]. This ensures that TCC calculations may always be performed with polynomial, as opposed to factorial, complexity. Furthermore, TCC is rigorously extensive, exactly equivalent to FCI for a complete cluster operator, $\hat{T}$, and invariant to rotations in the underlying orbital spaces, $\{\psi_i\}$ and $\{\psi_a\}$ However, the non-Hermitian nature of $\bar{H}$,

$$\left(e^{-\hat{T}}\hat{H}e^{\hat{T}}\right)^\dagger = e^{\hat{T}^\dagger}\hat{H}e^{-\hat{T}^\dagger} \neq e^{-\hat{T}}\hat{H}e^{\hat{T}} \tag{2.81}$$

coupled with the projective determination of the equations to be solved, effectively eliminates the property of a variational upper bound from the TCC approach, and also complicates the calculation of both ground- and excited-state properties.

An alternative scheme to the projective approach of TCC is to instead insert the exponential wavefunction ansatz into the quantum mechanical energy

expecation value expression,

$$\frac{\langle\Phi|\hat{H}|\Phi\rangle}{\langle\Phi|\Phi\rangle} = \frac{\langle\Phi_0|e^{\hat{T}^\dagger}\hat{H}e^{\hat{T}}|\Phi_0\rangle}{\langle\Phi_0|e^{\hat{T}^\dagger}e^{\hat{T}}|\Phi_0\rangle} \tag{2.82}$$

and to minimize this functional with respect to the set of cluster amplitudes, $\{T_a^i\} \cup \{T_{ab}^{ij}\} \cup \ldots$, in order to obtain a prediction of the ground-state energy. This is Variational Coupled Cluster[51] (VCC).

The VCC energy functional may be written in an explicitly linked form,

$$E_{\mathrm{VCC}} = \frac{\langle\Phi_0|e^{\hat{T}^\dagger}\hat{H}e^{\hat{T}}|\Phi_0\rangle}{\langle\Phi_0|e^{\hat{T}^\dagger}e^{\hat{T}}|\Phi_0\rangle} = \langle\Phi_0|e^{\hat{T}^\dagger}\hat{H}e^{\hat{T}}|\Phi_0\rangle_L, \tag{2.83}$$

since the unlinked terms present in the numerator and denominator exactly cancel, independent of the truncation of $\hat{T}$. Take, for example, the case of VCCD ($\hat{T} = \hat{T}_2$),

$$\begin{aligned}
\frac{\langle e^{\hat{T}_2^\dagger}\hat{H}e^{\hat{T}_2}\rangle}{\langle e^{\hat{T}_2^\dagger}e^{\hat{T}_2}\rangle} &= \frac{\langle\hat{H}\rangle + 2\langle\hat{H}\hat{T}_2\rangle + \langle\hat{T}_2^\dagger\hat{H}\hat{T}_2\rangle + \langle\hat{T}_2^\dagger\hat{H}\hat{T}_2\rangle + \ldots}{1 + \langle\hat{T}_2^\dagger\hat{T}_2\rangle + \ldots} \\
&= \frac{\begin{array}{c}\langle\hat{H}\rangle_L + 2\langle\hat{H}\hat{T}_2\rangle_L + \langle\hat{T}_2^\dagger\hat{H}\hat{T}_2\rangle_L + \langle\hat{H}\rangle\langle\hat{T}_2^\dagger\hat{T}_2\rangle \\ + \langle\hat{T}_2^\dagger\hat{H}\hat{T}_2^2\rangle_L + 2\langle\hat{H}\hat{T}_2\rangle\langle\hat{T}_2^\dagger\hat{T}_2\rangle + \ldots\end{array}}{1 + \langle\hat{T}_2^\dagger\hat{T}_2\rangle + \ldots} \\
&= \frac{\left(\langle\hat{H}\rangle_L + 2\langle\hat{H}\hat{T}_2\rangle_L + \langle\hat{T}_2\hat{H}\hat{T}_2\rangle_L + \ldots\right)\left(1 + \langle\hat{T}_2^\dagger\hat{T}_2\rangle + \ldots\right)}{1 + \langle\hat{T}_2^\dagger\hat{T}_2\rangle + \ldots} \\
&= \langle\hat{H}\rangle_L + 2\langle\hat{H}\hat{T}_2\rangle_L + \langle\hat{T}_2\hat{H}\hat{T}_2\rangle_L + \ldots \\
&= \langle e^{\hat{T}_2^\dagger}\hat{H}e^{\hat{T}_2}\rangle_L.
\end{aligned} \tag{2.84}$$

VCC is, like TCC, therefore rigorously extensive. It is noteworthy, however, that this cancellation, leaves uncancelled EPV terms present, making $\langle e^{\hat{T}_2^\dagger}\hat{H}e^{\hat{T}_2}\rangle_L$ an infinite expression; it does not terminate, even when the excitation rank exceeds the number of electrons. VCC is also exactly equivalent to FCI for a complete cluster operator, and invariant to rotations in the underlying orbital spaces. However, in contrast to the Traditional Coupled Cluster method outlined above, the operator $e^{\hat{T}^\dagger}\hat{H}e^{\hat{T}}$ is Hermitian,

$$(e^{\hat{T}^\dagger}\hat{H}e^{\hat{T}})^\dagger = (e^{\hat{T}})^\dagger\hat{H}^\dagger(e^{\hat{T}^\dagger})^\dagger = e^{\hat{T}^\dagger}\hat{H}e^{\hat{T}} \tag{2.85}$$

and, even for an incomplete cluster operator, a calculated VCC energy is an upper bound on the exact ground-state Schrödinger eigenvalue. VCC therefore represents the perfect quantum-chemical method, possessing all of the methodological

properties widely regarded to be important. Unfortunately, however, there is no known analogue of the Campbell-Baker-Hausdorff expansion within VCC theory, which renders the method factorially expensive, even with a cluster operator restricted to single and double excitations of the reference determinant, severely limiting its application only to extremely simple chemical systems with small basis sets.

There are, in fact, many more 'alternative' Coupled Cluster methods, such as Unitary Coupled Cluster[46–50] (UCC), Improved Coupled Cluster[52] (ICC), Extended Coupled Cluster[53] (ECC) and Quadratic Coupled Cluster[55] (QCC). In fact, each of these approaches may be related to the VCC method, or viewed as an approximation to it. The UCC method, in particular, defines the anti-Hermitian operator,

$$\hat{\sigma} = \hat{T} - \hat{T}^{\dagger}, \tag{2.86}$$

from which an exponential ansatz,

$$|\Phi\rangle = e^{\hat{\sigma}}|\Phi_0\rangle, \tag{2.87}$$

is constructed. Insertion into the quantum mechanical energy expectation value expression, in the same manner as the exponential ansatz in VCC theory, yields the following functional,

$$E_{\mathrm{UCC}} = \langle e^{\hat{\sigma}^{\dagger}} \hat{H} e^{\hat{\sigma}} \rangle \tag{2.88}$$

which, again, should be minimized with respect to the cluster amplitudes in order to obtain a ground-state energy. The denominator of the above expression is automatically unity, due to the unitary nature of $e^{\hat{\sigma}}$. The UCC energy is an infinite expression, much like $\langle e^{\hat{T}^{\dagger}} \hat{H} e^{\hat{T}} \rangle_L$, due to the presence of both the excitation operator, $\hat{T}$, and the de-excitation operator, $\hat{T}^{\dagger}$, in the argument of the exponential, although it has been shown that the series converges rapidly[47]. At low orders, the UCC terms generated are the same as those present in VCC, but higher-order contributions become significantly more complicated than the equivalent VCC terms[50], and, for this reason, UCC will be given no further theoretical consideration in this thesis.

The existence of these alternative schemes motivates their numerical benchmarking, with the goal of determining whether they are superior to the TCC method. Numerous such studies have already been undertaken, with the unanimous conclusion that the alternative Coupled Cluster ansätze are vastly superior to TCC[28, 41–45]. Of particular importance to this thesis is the VCC method, which has been demonstrated to remain particularly robust and accurate when non-dynamic correlation becomes strong, such as when modelling the dissociation of multiply-bonded molecules. It is widely known that the CCSD method perform quite poorly in these situations, typically predicting erroneous maxima in potential energy curves at long bond lengths, suggesting an unphysical long-range repulsion. When perturbative corrections, to be discussed in more detail in Chapter 7, such as (T) are applied to this poor starting point, the predicted potential energy curves can become catastrophically wrong, although more advanced perturbative corrections are in development[80–85]. Calculated VCCSD energies, on the other hand, are guaranteed by the variational upper bound property to never drop below the exact ground-state Schrödinger eigenvalue, which means, in practice, a VCCSD potential energy curve will remain above the corresponding FCI curve, and not predict the same unphysical effects problematic for CCSD. An illustration of the difference between CCSD and VCCSD relative to FCI is given in Figure 2.1, for the example of breaking the triple bond in acetylene. These calculations have been performed with the Molpro package of *ab initio* quantum chemistry programs[86, 87].

In fact, it is well-known that in order to achieve a physically correct description of the dissociation of strongly-correlated molecules such as dinitrogen, N_2, a TCC scheme must include not just single and double, or even triple, but a full treatment of quadruple excitations, CCSDTQ. A VCC scheme, however, displays an interesting asymmetry with TCC in this regard, requiring only single and double excitations to achieve the same qualitative effect. Analyses, such as those of Kutzelnigg[13, 54], elucidate this distinction by noting that the TCC-based CCD method is deficient in terms present in VCCD, beginning at fourth-order in the cluster amplitudes, $\mathcal{O}(T^4)$, and that this deficiency remains in both CCSD and CCSDT. It is not until quadruple excitations are added, CCSDTQ, that TCC

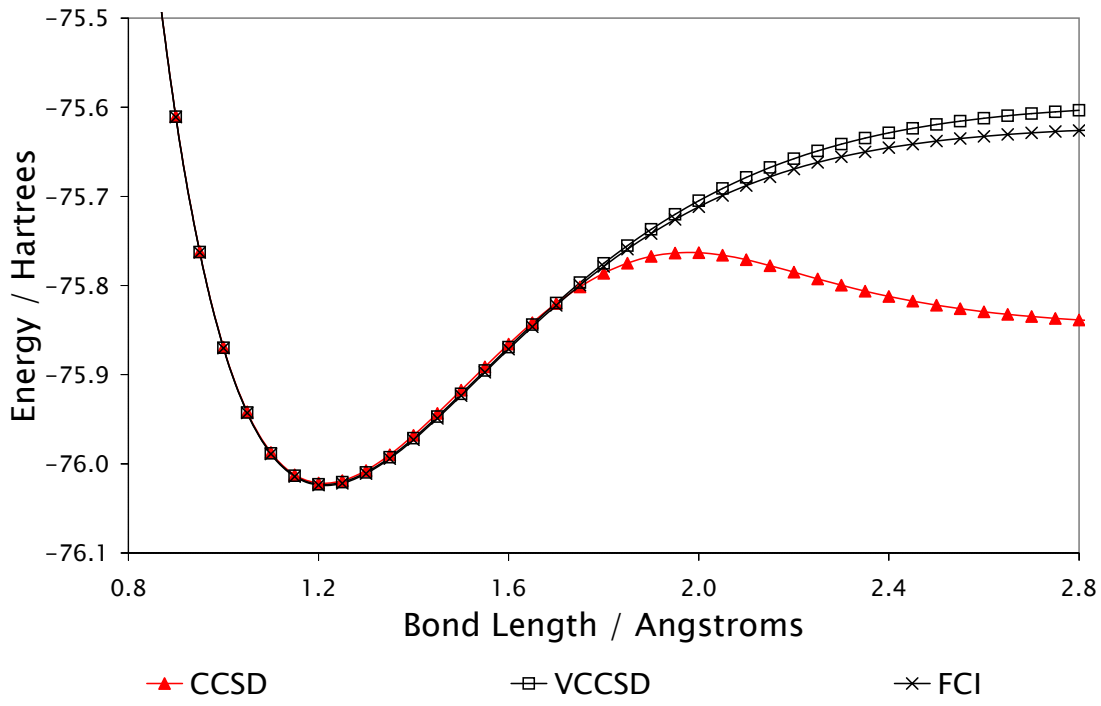


Figure 2.1: Calculated potential energy curves for the stretching of the carbon-carbon triple bond in acetylene, C_2H_2, with the C-H bond length fixed at 1.06Å, and with the STO-3G basis set.

converges closer to VCC; they are, of course, identical and equivalent to FCI in the limit of a complete cluster operator.

The conclusion is that the poor performance of the TCC methods to describe, for example, bond breaking, is not the fault of the CC wavefunction, but is instead associated with the non-Hermitian similarity-transformed Hamiltonian and projective determination of the equations to be solved. It is unfortunate, then, that the recent study of Evangelista[56] has demonstrated the unfavourable computational scaling of the alternative Coupled Cluster methods to significantly outweigh any potential gains in quantitative accuracy. This is true even of the QCCSD method, possibly the simplest correction of CCSD towards VCCSD, which scales as $\mathcal{O}(v^6)$, much more expensive than the limiting step of a CCSD calculation, $\mathcal{O}(o^2v^4)$.

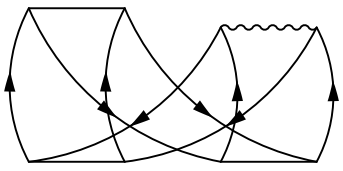
A superior CC-like method for which energies are calculated through the minimization of a functional, and which yields at least an approximately-fulfilled variational upper bound on the ground-state Schrödinger eigenvalue while never exceeding the CCSD-like computational complexity of $\mathcal{O}(o^2v^4)$, would be highly

desirable. The construction of such a method is the subject of this thesis.

2.9 Appendix: Diagrammatic Notation in Coupled Cluster Theory

Throughout this thesis, the terms present in Coupled Cluster expressions, particularly those of Variational Coupled Cluster, may be represented diagrammatically for convenience. There is a one-to-one correspondence between a topologically unique Coupled Cluster diagram and a corresponding algebraic expression, and a set of rules exist for interpreting such diagrams. For example, in these diagrams, lines going down represent occupied orbitals, and lines going up represent virtual orbitals, and should be labelled with the indices $\{i, j, k, \ldots\}$ and $\{a, b, c, \ldots\}$ respectively. Horizontal lines represent cluster amplitudes, and wavy lines represent two-electron integrals, with the indices corresponding to the order in which the occupied and virtual lines are attached to each vertex (an amplitude or integral). The sign of the term equals -1 raised to the power of the sum of the number of occupied lines plus the number of loops. The numerical prefactor of each term equals $1/2$ raised to the power of the number of equivalent line pairs (a pair of occupied or vitual lines is equivalent if they begin and end on the same vertex) plus the number of equivalent amplitude vertex pairs (a pair of amplitude vertices is equivalent if they are connected to the Hamiltonian in an identical way). There are additional rules for open diagrams (where lines do not terminate on a vertex), which correspond to contributions to CC amplitude equations, however, the above rules are sufficient for the closed diagrams (energy contributions) encountered in this thesis.

The above rules are best understood through an example. Consider the following term, an $\mathcal{O}(T^3)$ contribution to VCCD,



for which there are four pairs of equivalent lines, no pairs of equivalent vertices, four occupied lines and two loops, and the diagram therefore represents the term

$$(-1)^{4+2}\left(\frac{1}{2}\right)^{4+0} T_{kl}^{ab}\, T_{cd}^{kl}\, T_{ij}^{cd}\, \langle ij||ab\rangle. \tag{2.89}$$

For more thorough introductions to Coupled Cluster diagrams, the reader is referred to the excellent reviews of Crawford and Schaefer[79] and of Bartlett and Musial[88].

Chapter 3

The Linked Pair Functional I: Fundamental Theory[†]

In the previous chapter, it was described how the failure of the CCSD method, based on Traditional Coupled Cluster theory, to adequately describe problems in which the single-determinantal reference wavefunction approximation of Hartree-Fock theory breaks down, is not necessarily the fault of the Coupled Cluster exponential ansatz. Rather, it is the lack of a variational upper bound property, caused by the non-Hermitian similarity-transformed Hamiltonian and the subsequent projective determination of the working equations, that allows TCC energies to fall unphysically below FCI energies. In contrast, the Variational Coupled Cluster method, which preserves the property that a calculated energy is a rigorous upper bound on the exact ground-state Schrödinger energy eigenvalue, performs significantly better. Since the factorial computational complexity of VCC, even at the doubles level, limits its application, it is natural to seek approximation schemes, but, historically, the approximation of even VCCD has been troublesome, with each of the proposed approaches flawed in some way. This chapter discusses these approaches, and further establishes a new quantum-chemical method, the Linked Pair Functional, with the potential to resolve this problem.

[†]Relevant publication:

[89] P. J. Knowles and B. Cooper, J. Chem. Phys. **133**, 224106 (2010).

3.1 Approximating VCC

As has been discussed above, Coupled Cluster theory differs from, for example, Configuration Interaction, by the ansatz of an exponential parameterisation of the wavefunction, modelled as the action of an exponential cluster operator on a single-determinantal reference wavefunction. As mentioned previously, the action of this exponential operator on the reference determinant can be understood through the Maclaurin series of the exponential function,

$$e^{\hat{T}}|\Phi_0\rangle = \sum_{n=0}^{\infty} \frac{\hat{T}^n}{n!}|\Phi_0\rangle. \tag{3.1}$$

The VCC functional, written with an uncancelled denominator, and stated again here for convenience as,

$$E_{\mathrm{VCC}} = \frac{\langle e^{\hat{T}^\dagger}\hat{H}e^{\hat{T}}\rangle}{\langle e^{\hat{T}^\dagger}e^{\hat{T}}\rangle} = \langle e^{\hat{T}^\dagger}\hat{H}e^{\hat{T}}\rangle_L. \tag{3.2}$$

therefore possesses non-zero contributions involving excitations up to the number of electrons. When the denominator is cancelled with the unlinked parts of the numerator to leave an explicitly linked functional, uncancelled Exclusion Principle Violating (EPV) terms remain, such that the expression does not terminate, and is therefore infinite. In contrast to TCC, for which the Campbell-Baker-Hausdorff expansion[27] applies, there is no simplification of the VCC functional, and it therefore has factorial computational complexity.

In order to arrive at a computationally tractable approximation to the VCC functional, the natural first attempt is to truncate the functional in some way, typically by direct truncation of the exponential operator itself. However, this is problematic, and an important objection to this truncation is the convergence of its series. Although the exponential function converges for all real (and complex) values of its argument[90, 91], the convergence is not rapid in general as is illustrated in Table 3.1, except for very small values of the argument. Besides the obvious problem of attempting to ascribe a measure of size, not to a number, but to the operator, $\hat{T}$, it cannot be true that $\hat{T}$ is 'small' for *all* systems of chemical interest. This renders a low-order truncation of the series inappropriate since there is then no guarantee that $1+\hat{T}$, for example, will be a good approximation to $e^{\hat{T}}$ in general.

Table 3.1: Convergence of the Maclaurin series of e^x for a selection of x values.

N	$\sum_{n=0}^{N} \frac{0.5^n}{n!}$	Relative Error	$\sum_{n=0}^{N} \frac{1.0^n}{n!}$	Relative Error	$\sum_{n=0}^{N} \frac{2.0^n}{n!}$	Relative Error
0	1.00000	-3.93×10^{-1}	1.00000	-6.32×10^{-1}	1.00000	-8.65×10^{-1}
1	1.50000	-9.02×10^{-2}	2.00000	-2.64×10^{-1}	3.00000	-5.94×10^{-1}
2	1.62500	-1.44×10^{-2}	2.50000	-8.03×10^{-2}	5.00000	-3.23×10^{-1}
3	1.64583	-1.75×10^{-3}	2.66667	-1.90×10^{-2}	6.33333	-1.43×10^{-1}
4	1.64844	-1.72×10^{-4}	2.70833	-3.66×10^{-3}	7.00000	-5.27×10^{-2}
5	1.64870	-1.42×10^{-5}	2.71667	-5.94×10^{-4}	7.26667	-1.66×10^{-2}
6	1.64872	-1.00×10^{-6}	2.71806	-8.32×10^{-5}	7.35556	-4.53×10^{-3}
7	1.64872	-6.22×10^{-8}	2.71825	-1.02×10^{-5}	7.38095	-1.10×10^{-3}
8	1.64872	-3.44×10^{-9}	2.71828	-1.13×10^{-6}	7.38730	-2.37×10^{-4}
9	1.64872	-1.71×10^{-10}	2.71828	-1.11×10^{-7}	7.38871	-4.65×10^{-5}
10	1.64872	-7.74×10^{-12}	2.71828	-1.00×10^{-8}	7.38899	-8.31×10^{-6}

However, there is an even more serious flaw in this scheme of approximation. Consider, for example, truncating the exponential to be linear in the cluster operator; $e^{\hat{T}} \rightarrow 1 + \hat{T}$. Inserting this approximation into the first form of the VCC functional yields Variational CI,

$$E_{\mathrm{VCC}} = \frac{\langle\Phi_0|e^{\hat{T}^\dagger}\hat{H}e^{\hat{T}}|\Phi_0\rangle}{\langle\Phi_0|e^{\hat{T}^\dagger}e^{\hat{T}}|\Phi_0\rangle} \longrightarrow \frac{\langle\Phi_0|(1+\hat{T})^\dagger\hat{H}(1+\hat{T})|\Phi_0\rangle}{\langle\Phi_0|(1+\hat{T})^\dagger(1+\hat{T})|\Phi_0\rangle} = E_{CI}, \tag{3.3}$$

which is not extensive. Alternatively, insertion into the second, explicitly linked form of the VCC functional yields CEPA(0),

$$E_{\mathrm{VCC}} = \langle\Phi_0|e^{\hat{T}^\dagger}\hat{H}e^{\hat{T}}|\Phi_0\rangle_L \longrightarrow \langle\Phi_0|(1+\hat{T})^\dagger\hat{H}(1+\hat{T})|\Phi_0\rangle_L = E_{\mathrm{CEPA(0)}}, \tag{3.4}$$

which is not exact, even when the cluster operator is complete. Thus, truncation of the exponential operator has, in each case, eliminated at least one of the important methodological properties possessed by a true Coupled Cluster method. This is, in fact, indicative of the more general case; in the first form of the functional, truncation of the exponential operator to *any* polynomial degree destroys the exact cancellation of the VCC denominator with the unlinked parts of the VCC numerator, leaving uncancelled unlinked terms that lead to unphysical scaling of the energy with system size and thus the loss of extensivity. If the exact cancellation of the denominator is performed first, uncancelled Exclusion Principle Violating (EPV) terms remain such that $\langle e^{\hat{T}^\dagger}\hat{H}e^{\hat{T}}\rangle_L$ does not terminate, as discussed above. Any finite approximation to this infinite expression, such as

that generated by truncation of the exponential operator to any finite polynomial degree, is thus necessarily incomplete and incorrect, even when the cluster operator is itself complete.

An alternative to the direct approximation of the VCC functional is to correct upon a TCC starting point. As such, asymmetric expectation value expressions have been proposed, such as Improved Coupled Cluster[52] (ICC) and Extended Coupled Cluster[53] (ECC), which suggest hierarchies of methods stepping systematically from TCC to VCC. They have the disadvantage, however, that there is typically no rigorous guarantee of extensivity except at the extremes of the hierarchies. The simplest possible correction of TCC towards VCC is captured by the Quadratic Coupled Cluster[55] (QCC) method. Unfortunately, in light of the analyses of Kutzelnigg[54] and others, the lowest-order correction terms are inevitably $\mathcal{O}(T^4)$, and contain contributions to the energy that cannot be computed in less than $\mathcal{O}(v^6)$ time, which is significantly more expensive than the limiting $\mathcal{O}(o^2v^4)$ step in a CCSD calculation, since, typically, $v \gg o$. For this reason, the use of such correction schemes have not found widespread application within the quantum-chemical community.

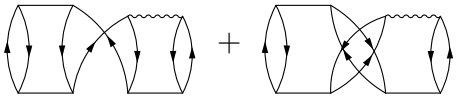
3.2 The Internal Mathematical Structure of Variational Coupled Cluster Doubles

In summary of the discussions of the previous section, the construction of approximations to the Variational Coupled Cluster method is quite problematic. Direct approximation schemes relying on the truncation of the exponential operator, $e^{\hat{T}}$, to, for example, $1+\hat{T}$ reduce the VCC functional only to that of Variational CI or CEPA(0). A more relaxed truncation, perhaps to $1+\hat{T}+\frac{1}{2!}\hat{T}^2$, does not resolve the problem and still results in the loss of extensivity or an exact treatment of limiting systems. Furthermore, the cost of computing terms such as $\left(\frac{1}{2!}\right)^2 \langle(\hat{T}_2^\dagger)^2 \hat{H} \hat{T}_2^2\rangle_L$ is, at $\mathcal{O}(v^6)$ complexity, significantly higher than even the limiting step in a CCSD calculation. Attempts to correct, for example, CCSD towards VCCSD suffer from this same unsatisfactorarily high computational scaling.

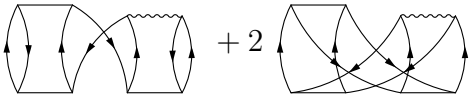
It appears, however, that few authors have recognized the VCC functional to possess an internal mathematical structure, and that this structure can be exploited. For simplicity, consider the VCC functional for the case of a cluster operator restricted to contain only double excitations of the single-determinantal reference wavefunction, $\hat{T} = \hat{T}_2$, which defines the VCCD method,

$$E_{\mathrm{VCCD}} = \frac{\langle e^{\hat{T}_2^\dagger} \hat{H} e^{\hat{T}_2} \rangle}{\langle e^{\hat{T}_2^\dagger} e^{\hat{T}_2} \rangle} = \langle e^{\hat{T}_2^\dagger} \hat{H} e^{\hat{T}_2} \rangle_L. \tag{3.5}$$

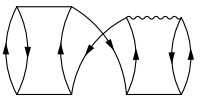
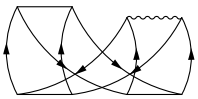
The term $\frac{1}{2!}\langle \hat{T}_2^\dagger \hat{H} \hat{T}_2^2 \rangle_L$ is the lowest-order term present in VCCD that is omitted by CEPA(0). The contributions to this term, represented diagrammatically, are given in Fig. 3.1. These terms, labelled $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$, obey certain relationships for the case of a 2-electron system. In particular, the terms $\mathcal{A}$ and $\mathcal{D}$ cancel,

$\mathcal{A} + \mathcal{D} =$  $= 0,$ (3.6)

and the remaining terms satisfy,

$\mathcal{B} + 2\,\mathcal{C} =$  $= 0,$ (3.7)

in this limit. Therefore, in this case of two electrons, the complete VCCD $\mathcal{O}(T^3)$ contribution, $\mathcal{A} + \mathcal{B} + \mathcal{C} + \mathcal{D}$, may be captured in infinitely many ways,

$$\begin{aligned} \mathcal{A} + \mathcal{B} + \mathcal{C} + \mathcal{D} &= \mathcal{B} + \mathcal{C} \\ &= \frac{1}{2}(1-\lambda)\mathcal{B} - \lambda\,\mathcal{C} \\ &= \frac{1}{2}(1-\lambda)\ \text{  }\ - \lambda\ \text{  }\ , \end{aligned} \tag{3.8}$$

corresponding to the continuously adjustable weightings of the $\mathcal{B}$ and $\mathcal{C}$ terms, controlled by the parameter λ. Similar relationships have, in fact, been noted by Huntington and Nooijen[74] also to occur in the TCCSD residual. Three values of the parameter λ, in particular, stand out. For $\lambda = 0$, the $\mathcal{C}$ term vanishes completely. Similarly, for $\lambda = +1$, the $\mathcal{B}$ term is switched off entirely. The case of $\mathcal{B} + \mathcal{C}$, for which the weightings of the $\mathcal{B}$ and $\mathcal{C}$ terms are equal to the true weightings as the terms appear in VCCD, corresponds to the case $\lambda = -1$ with this parameterisation. Similar cancellations occur between terms throughout the VCCD functional, through all orders of the cluster amplitudes.

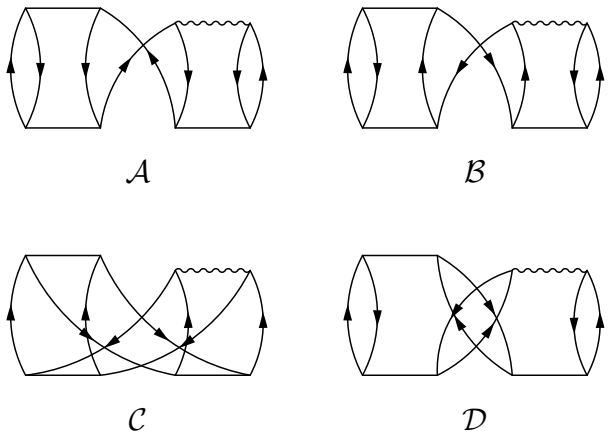


Figure 3.1: Linked $\mathcal{O}(T^3)$ terms contributing to VCCD.

While it remains true that any finite truncation of $\langle e^{\hat{T}_2^\dagger}\hat{H}e^{\hat{T}_2}\rangle_L$ will not be exact, this mathematical structure implies that it is possible to construct an approximation to this form of the VCC functional, and, crucially, one that remains exact for 2 electrons, by capturing only a subset of the contributing terms, albeit to infinite order. The Linked Pair Functional is a quantum-chemical method that exploits this property. It is first introduced and motivated from the more well-known problem[8, 9] of attempting to construct a post-Hartree-Fock Hermitian energy functional with the properties of rigorous extensivity, orbital invariance and an exact treatment of 2-electron systems before its relationship to VCCD is discussed. It is particularly satisfying that a potential solution to this problem is one that also solves the problem of constructing an approximation to VCCD that possesses these same properties.

3.3 Linked Pair Functional Theory

Consider the Configuration Interaction Doubles functional,

$$E_{\mathrm{CID}} = \frac{\langle\Phi_0|(1+\hat{T}_2)^\dagger\hat{H}(1+\hat{T}_2)|\Phi_0\rangle}{\langle\Phi_0|(1+\hat{T}_2)^\dagger(1+\hat{T}_2)|\Phi_0\rangle}, \tag{3.9}$$

and the CEPA(0) Doubles functional,

$$E_{\mathrm{CEPA(0)}} = \langle\Phi_0|(1+\hat{T}_2)^\dagger\hat{H}(1+\hat{T}_2)|\Phi_0\rangle_L. \tag{3.10}$$

Disregarding the effects of single excitations, the CID functional is equivalent to FCI for a limiting system containing only two electrons, but is not extensive. The CEPA(0) functional, on the other hand, is rigorously extensive since it contains only linked terms, but is not equivalent to CID for a 2-electron system. However, by partitioning these functionals into contributions to the reference and correlation energies,

$$\begin{aligned} E_{\mathrm{CID}} &= \frac{\langle\Phi_0|(1+\hat{T}_2)^\dagger \hat{H}(1+\hat{T}_2)|\Phi_0\rangle}{\langle\Phi_0|(1+\hat{T}_2)^\dagger(1+\hat{T}_2)|\Phi_0\rangle} \\ &= \frac{\langle\hat{H}\rangle + 2\langle\hat{H}\,\hat{T}_2\rangle + \langle\hat{T}_2^\dagger\,\hat{H}\,\hat{T}_2\rangle}{1+\langle\hat{T}_2^\dagger\,\hat{T}_2\rangle} \\ &= \langle\hat{H}\rangle + \frac{2\langle\hat{H}\,\hat{T}_2\rangle + \langle\hat{T}_2^\dagger(\hat{H}-\langle\hat{H}\rangle)\hat{T}_2\rangle}{1+\langle\hat{T}_2^\dagger\,\hat{T}_2\rangle}, \end{aligned} \tag{3.11}$$

for the CID functional, and,

$$\begin{aligned} E_{\mathrm{CEPA(0)}} &= \langle\Phi_0|(1+\hat{T}_2)^\dagger \hat{H}(1+\hat{T}_2)|\Phi_0\rangle_L \\ &= \langle\hat{H}\rangle + 2\langle\hat{H}\,\hat{T}_2\rangle + \langle\hat{T}_2^\dagger\,\hat{H}\,\hat{T}_2\rangle_L \\ &= \langle\hat{H}\rangle + 2\langle\hat{H}\,\hat{T}_2\rangle + \langle\hat{T}_2^\dagger(\hat{H}-\langle\hat{H}\rangle)\hat{T}_2\rangle, \end{aligned} \tag{3.12}$$

for the CEPA(0) functional, it is apparent that they are remarkably similar. In fact, the numerator of the CID correlation energy is simply the CEPA(0) correlation energy. It also makes clear that the CID numerator is itself extensive (since it contains only the fully-linked CEPA(0) terms), and the unlinked terms that violate extensivity enter only through division by the CID denominator; the CID numerator necessarily already scales physically with the system size since its fully linked nature makes it rigorously extensive, and any dependence of the denominator on the size of the system therefore disrupts this behaviour. Unfortunately, it is easy to see that the CID denominator grows with system size;

$$\langle\hat{T}_2^\dagger\hat{T}_2\rangle = \frac{1}{4}T_{ab}^{ij}T_{ij}^{ab} = \frac{1}{4}\sum_{i,j}^{N_{\mathrm{elec}}}\sum_{a,b}^{N_{\mathrm{virt}}}(T_{ij}^{ab})^2 \leq \frac{1}{4}\sum_{i,j}^{N'_{\mathrm{elec}}}\sum_{a,b}^{N'_{\mathrm{virt}}}(T_{ij}^{ab})^2 \tag{3.13}$$

if $N'_{\mathrm{elec}} \geq N_{\mathrm{elec}}$ and $N'_{\mathrm{virt}} \geq N_{\mathrm{virt}}$. Alternatively, the loss of extensivity caused by division by the CID denominator can be understood through the introduction of unlinked terms via the binomial theorem,

$$\begin{aligned} E_{\mathrm{CID}} &= \langle\hat{H}\rangle + \frac{2\langle\hat{H}\,\hat{T}_2\rangle + \langle\hat{T}_2^\dagger(\hat{H}-\langle\hat{H}\rangle)\hat{T}_2\rangle}{1+\langle\hat{T}_2^\dagger\,\hat{T}_2\rangle} \\ &= \langle\hat{H}\rangle + \Big(2\langle\hat{H}\,\hat{T}_2\rangle + \langle\hat{T}_2^\dagger(\hat{H}-\langle\hat{H}\rangle)\hat{T}_2\rangle\Big)\Big(1-\langle\hat{T}_2^\dagger\hat{T}\rangle + \ldots\Big) \end{aligned} \tag{3.14}$$

This suggests that a functional that is potentially both exact and extensive could be achieved by simply dividing the CEPA(0) functional by a denominator that does not scale with system size. This has, of course, been attempted previously; for example, the Coupled Pair Functional[8] (CPF) of Ahlrichs et al. and Kollmar's functionals[72, 73] each attempt schemes in which the terms present in the numerator are scaled only by contributions to the denominator to which they are local, the Averaged Coupled Pair Functional[31] (ACPF) simply divides the term $\langle \hat{T}_2^\dagger \hat{T}_2 \rangle$ by the number of electrons, and alternative CEPA approaches attempt to eliminate the unphysical unlinked contributions from the working equations in order to achieve a similar effect[92]. These functional modifications, however, are typically either only approximately extensive, only approximately exact for 2 electrons, or not invariant to orbital rotations.

The Linked Pair Functional (LPF), however, posits that the effect of the CID denominator can be incorporated directly into the cluster amplitudes by an appropriate matrix transformation. The advantage of this scheme is that such constructions are fully linked tensor expressions, and an energy can be constructed that is therefore not only rigorously extensive, but also explicitly scalar and thus invariant to rotations in the underlying orbital spaces. Furthermore, an appropriate choice of the form of the transformation and its associated transformation matrix can render the approach additionally exactly equivalent to CID for a 2-electron system. The explicit form of Linked Pair Functional Doubles (LPFD) theory is given below.

A ground-state LPFD energy is calculated as the minimum of the LPFD energy functional,

$$E_{\mathrm{LPFD}} = \langle \hat{H} \rangle + 2\langle \hat{H}\, {}_2\hat{T} \rangle + \langle {}_1\hat{T}^\dagger (\hat{H} - \langle \hat{H} \rangle)\, {}_1\hat{T} \rangle, \tag{3.15}$$

which takes the same form as the CEPA(0) functional except that left subscripts have been inserted on the cluster operators; ${}_2\hat{T}$ and ${}_1\hat{T}$. These are doubles-only excitation operators,

$${}_q\hat{T}|\Phi_0\rangle = \frac{1}{4} \sum_{i,j,a,b} {}_qT_{ab}^{ij} b^\dagger j a^\dagger i |\Phi_0\rangle = \frac{1}{4}\, {}_qT_{ab}^{ij} |\Phi_{ij}^{ab}\rangle, \tag{3.16}$$

where the coefficients, $\{{}_qT_{ab}^{ij}\}$, are the transformed cluster amplitudes. These are defined as the contraction of a transforming tensor, $\mathbf{U}^{-\frac{q}{2}}$, with the untransformed

cluster amplitudes,

$$ {}_qT_{ab}^{ij} = (\mathbf{U}^{-\frac{q}{2}}\mathbf{T})_{ab}^{ij} = \frac{1}{2}(\mathbf{U}^{-\frac{q}{2}})_{kl}^{ij}\, T_{ab}^{kl}. \tag{3.17} $$

Since it is required that the transformed cluster amplitudes inherit the fermionic antisymmetry properties of the untransformed cluster amplitudes, it is necessary that U_{kl}^{ij} also possesses these properties,

$$ U_{kl}^{ji} = -U_{kl}^{ij}, \tag{3.18} $$

$$ U_{lk}^{ij} = -U_{kl}^{ij}. \tag{3.19} $$

Thus, there is a one-to-one relationship between the unique elements of the 4-component tensor, U_{kl}^{ij}, and the elements of the two-index matrix, $\mathbf{U}$, with rows and columns labelled by the unique electron pairs, for which the composite indices $ij, kl \in \{(m,n),\ m > n\}$ stand. With this alternative representation, it makes sense to talk of a transformation matrix and also of its powers, which can be defined simply as matrix powers. This gives meaning to $\mathbf{U}^{-\frac{q}{2}}$, used above. The definition of $\mathbf{U}^2$, for example, is as follows,

$$ (\mathbf{U}^2)_{kl}^{ij} = (\mathbf{U}\mathbf{U})_{kl}^{ij} = \sum_{m>n} U_{mn}^{ij} U_{kl}^{mn} = \frac{1}{2} U_{mn}^{ij} U_{kl}^{mn}, \tag{3.20} $$

and more general powers may be computed first by diagonalizing the $N_{\text{pair}} \times N_{\text{pair}}$ representation of $\mathbf{U}$, raising the resulting diagonal matrix to the appropriate power and then reversing the diagonalization.

The transformation matrix itself takes the following form,

$$ U_{kl}^{ij} = \delta_{kl}^{ij} + \Delta_{kl}^{ij} \tag{3.21} $$

$$ \Delta_{kl}^{ij} = \lambda\, \eta_{kl}^{ij} + \frac{1}{2}\,(1-\lambda)\,(1-\tau_{ij})\,(1-\tau_{kl})\,\delta_k^i\,\eta_l^j, \tag{3.22} $$

where τ_{ij} permutes the labels i, j in what follows, and where,

$$ \eta_j^i = \langle \hat{T}^\dagger j i^\dagger \hat{T} \rangle = \frac{1}{2} T_{ab}^{ik}\, T_{jk}^{ab} \tag{3.23} $$

$$ \eta_{kl}^{ij} = \langle \hat{T}^\dagger k l j^\dagger i^\dagger \hat{T} \rangle = \frac{1}{2} T_{ab}^{ij}\, T_{kl}^{ab} \tag{3.24} $$

are the one- and two-hole reduced density matrices[93]. It should be noted that this explicit form for U_{kl}^{ij}, satisfies the fermionic antisymmetry criteria outlined

above, as can be confirmed trivially from the definitions given. The quantity λ is a continuously-adjustable real scalar parameter.

It has already been mentioned that since the LPFD energy functional corresponds simply to the CEPA(0) functional in which a transformation of the cluster amplitudes is carried out, and that these new amplitudes are themselves fully linked tensors, the functional contains no unlinked terms, making it rigorously extensive. Also from the fact that the new amplitudes are fully linked tensors, and the terms contributing to the functional contain no unsummed indices, for example,

$$\langle \hat{H}\, {}_2\hat{T} \rangle = \frac{1}{4} \langle ij||ab \rangle\, {}_2T_{ij}^{ab}, \tag{3.25}$$

the energy is therefore a scalar, invariant to rotations in the underlying orbital spaces $\{\psi_i\}$ and $\{\psi_a\}$. Furthermore, consider the behaviour of the transformations in the 2-electron limit. In this limit, there is only a single unique electron pair, and the $\mathbf{U}$ matrix is thus simply a scalar. Therefore,

$$\begin{aligned} {}_qT_{ab}^{e\bar{e}} &= \frac{1}{2} (\mathbf{U}^{-\frac{q}{2}})_{kl}^{e\bar{e}} T_{ab}^{kl} \\ &= \frac{1}{2} (\mathbf{U}^{-\frac{q}{2}})_{e\bar{e}}^{e\bar{e}} T_{ab}^{e\bar{e}} + \frac{1}{2} (\mathbf{U}^{-\frac{q}{2}})_{\bar{e}e}^{e\bar{e}} T_{ab}^{\bar{e}e} \\ &= (\mathbf{U}^{-\frac{q}{2}})_{e\bar{e}}^{e\bar{e}} T_{ab}^{e\bar{e}} \\ &= \frac{T_{ab}^{e\bar{e}}}{(U_{e\bar{e}}^{e\bar{e}})^{\frac{q}{2}}}, \end{aligned} \tag{3.26}$$

which, for each term in the functional, introduces division by $U_{e\bar{e}}^{e\bar{e}}$, where e and $\bar{e}$ are labels for the two electrons present in the system, and for which the Einstein summation convention is not implied. The equivalent notation h and $\bar{h}$ will be used in later chapters when discussing 2-hole systems, or systems containing only 2 unoccupied (virtual) orbitals. From the definition of U_{kl}^{ij}, it further follows that $U_{e\bar{e}}^{e\bar{e}}$ equals the CID norm, since,

$$\langle \hat{T}_2^\dagger \hat{T}_2 \rangle = \frac{1}{4} T_{ab}^{ij} T_{ij}^{ab} = \frac{1}{4} T_{ab}^{e\bar{e}} T_{e\bar{e}}^{ab} + \frac{1}{4} T_{ab}^{\bar{e}e} T_{\bar{e}e}^{ab} = \frac{1}{2} T_{ab}^{e\bar{e}} T_{e\bar{e}}^{ab}, \tag{3.27}$$

and,

$$\eta_{e\bar{e}}^{e\bar{e}} = \eta_e^e = \eta_{\bar{e}}^{\bar{e}} = \frac{1}{2} T_{ab}^{e\bar{e}} T_{e\bar{e}}^{ab}, \tag{3.28}$$

and therefore,

$$\begin{aligned} U_{e\bar{e}}^{e\bar{e}} &= \delta_{e\bar{e}}^{e\bar{e}} + \lambda\,\eta_{e\bar{e}}^{e\bar{e}} + \frac{1}{2}(1-\lambda)\,(\delta_e^e\,\eta_{\bar{e}}^{\bar{e}} - \delta_{\bar{e}}^e\,\eta_e^{\bar{e}} - \delta_e^{\bar{e}}\,\eta_{\bar{e}}^e + \delta_{\bar{e}}^{\bar{e}}\,\eta_e^e) \\ &= 1 + \lambda\langle\hat{T}_2^\dagger\hat{T}_2\rangle + \frac{1}{2}(1-\lambda)\left(\langle\hat{T}_2^\dagger\hat{T}\rangle + \langle\hat{T}_2^\dagger\hat{T}\rangle\right) \\ &= 1 + \lambda\langle\hat{T}_2^\dagger\hat{T}_2\rangle + (1-\lambda)\langle\hat{T}_2^\dagger\hat{T}_2\rangle \\ &= 1 + (\lambda + 1 - \lambda)\langle\hat{T}_2^\dagger\hat{T}_2\rangle \\ &= 1 + \langle\hat{T}_2^\dagger\hat{T}_2\rangle. \end{aligned} \tag{3.29}$$

Notice the cancellation of the λ parameter. Thus, the LPFD amplitude transformations are exactly equivalent to direct division by the CID denominator for 2-electron systems. Hence, the LPFD functional reduces exactly, by construction, to the CID functional, and is therefore additionally exact in this limit, independent of λ. For more general systems, the effect of the matrix inverse is to introduce division by those parts of the CID denominator that are coupled to a given pair of orbitals, rather than by the complete denominator.

Since no specific value for the λ parameter can be assigned without first seeking additional information, LPFD methods with different values for this parameter will be discussed. To simplify such discussions, the nomenclature LPFD(λ) will be used for LPFD methods with specific values of λ. The purpose of the λ parameter, which simply controls the weighting of the one- and 2-hole density matrices contributing to the transformation matrix, will be explained further in the following section.

The LPFD scheme possesses an extremely theoretically aesthetic set of methodological properties, which, for completeness, are summarised below.

- The ground-state energy is calculated by variational minimization of a functional. The error in a calculated energy is thus second-order in any remaining errors in the cluster amplitude parameters.
- The functional contains fully linked terms only, and is therefore rigorously extensive.
- It is exactly equivalent to CID for a limiting 2-electron system.

- The energy is a scalar that is invariant to rotations in the underlying orbital spaces $\{\psi_i\}$ and $\{\psi_a\}$.
- The energy is not a true upper bound on the exact ground-state Schrödinger energy eigenvalue, but variational minimization of the functional gives a theory that satisfies the generalized Hellmann-Feynman theorem[67].
- The limiting computational complexity of the method is $\mathcal{O}(o^2v^4)$, the same as CCSD.

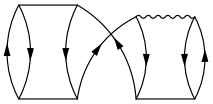
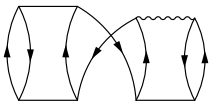
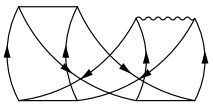
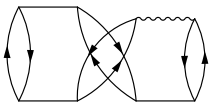
It therefore appears that the LPFD method, when combined with a suitable treatment of single excitations, outlined in the following chapter, resolves the long-standing problem[8, 9] of finding a computationally practical approximate solution to the Born-Oppenheimer electronic Schrödinger Equation that can be formulated through the minimization of a functional that is exact, extensive and invariant to orbital rotations. Furthermore, it can be shown that the LPFD method possesses a very deep connection with the VCCD method, which will be discussed in the following section.

3.4 Relationship of LPFD to VCCD

Further theoretical understanding of the workings of the LPFD method can be gained by an examination of the terms present in the transformed cluster amplitudes through the application of the binomial theorem to the matrix $\mathbf{U}^{-\frac{q}{2}}$.

$$\begin{aligned}
{}_qT_{ab}^{ij} &= \frac{1}{2}(\mathbf{U}^{-\frac{q}{2}})_{kl}^{ij}T_{ab}^{kl} \\
&= \frac{1}{2}((\mathbf{1}+\mathbf{\Delta})^{-\frac{q}{2}})_{kl}^{ij}T_{ab}^{kl} \\
&= \frac{1}{2}\sum_{n=0}^{\infty}\binom{-\frac{q}{2}}{n}(\mathbf{\Delta}^n)_{kl}^{ij}T_{ab}^{kl} \\
&= \frac{1}{2}(\mathbf{\Delta}^0)_{kl}^{ij}T_{ab}^{kl} - \frac{q}{4}(\mathbf{\Delta})_{kl}^{ij}T_{ab}^{kl} + \dots \\
&= \sum_{k>l}\delta_{kl}^{ij}T_{ab}^{kl} - \frac{q}{4}(\mathbf{\Delta})_{kl}^{ij}T_{ab}^{kl} + \dots \\
&= T_{ab}^{ij} - \frac{q}{4}(\mathbf{\Delta})_{kl}^{ij}T_{ab}^{kl} + \dots
\end{aligned} \tag{3.30}$$

Table 3.2: Linked $\mathcal{O}(T^3)$ contributions to VCCD, where $\langle \hat{T}^\dagger \hat{T} \rangle = \frac{1}{4} T_{ab}^{ij} T_{ij}^{ab}$, and where e and $\bar{e}$ label the two spinorbitals occupied in the reference wavefunction.

	$\frac{1}{2!}\langle \hat{T}_2^\dagger \hat{H} \hat{T}_2^2 \rangle_L$ Diagram	$\frac{1}{2!}\langle \hat{T}_2^\dagger \hat{H} \hat{T}_2^2 \rangle_L$ Term	2 Electrons
$\mathcal{A}$		$\frac{1}{4} T_{ij}^{ac} T_{cd}^{kl} T_{kl}^{db} \langle ij \vert\vert ab \rangle$	$T_{e\bar{e}}^{ac} T_{cd}^{e\bar{e}} T_{e\bar{e}}^{db} \langle e\bar{e} \vert\vert ab \rangle$
$\mathcal{B}$		$\frac{1}{4} T_{ik}^{ab} T_{cd}^{kl} T_{lj}^{cd} \langle ij \vert\vert ab \rangle$ $= -\frac{1}{2} T_{ik}^{ab} \eta_j^k \langle ij \vert\vert ab \rangle$	$-T_{e\bar{e}}^{ab} \langle \hat{T}^\dagger \hat{T} \rangle \langle e\bar{e} \vert\vert ab \rangle$
$\mathcal{C}$		$\frac{1}{16} T_{kl}^{ab} T_{cd}^{kl} T_{ij}^{cd} \langle ij \vert\vert ab \rangle$ $= \frac{1}{8} T_{kl}^{ab} \eta_{ij}^{kl} \langle ij \vert\vert ab \rangle$	$\frac{1}{2} T_{e\bar{e}}^{ab} \langle \hat{T}^\dagger \hat{T} \rangle \langle e\bar{e} \vert\vert ab \rangle$
$\mathcal{D}$		$\frac{1}{2} T_{ik}^{ac} T_{cd}^{kl} T_{lj}^{db} \langle ij \vert\vert ab \rangle$	$-T_{e\bar{e}}^{ac} T_{cd}^{e\bar{e}} T_{e\bar{e}}^{db} \langle e\bar{e} \vert\vert ab \rangle$

Noting that $\mathbf{\Delta}$ is $\mathcal{O}(T^2)$, the leading contribution to the transformed cluster amplitudes is therefore the untransformed amplitudes,

$$ {}_q T_{ab}^{ij} = T_{ab}^{ij} + \mathcal{O}(T^3). \tag{3.31} $$

Therefore, the leading contribution to the LPFD energy is the CEPA(0) energy,

$$ \begin{aligned} E_{\mathrm{LPFD}} &= \langle \hat{H} \rangle + 2 \langle \hat{H}\, {}_2\hat{T} \rangle + \langle {}_1\hat{T}^\dagger (\hat{H} - \langle \hat{H} \rangle)\, {}_1\hat{T} \rangle \\ &= \langle \hat{H} \rangle + 2 \langle \hat{H} \hat{T}_2 \rangle + \langle \hat{T}_2^\dagger (\hat{H} - \langle \hat{H} \rangle) \hat{T}_2 \rangle + \mathcal{O}(T^3) \\ &= E_{\mathrm{CEPA(0)}} + \mathcal{O}(T^3). \end{aligned} \tag{3.32} $$

As discussed above, CEPA(0) is itself a low-order approximation of VCCD, correct to $\mathcal{O}(T^2)$, and LPFD is therefore equivalent to VCCD to $\mathcal{O}(T^2)$ also,

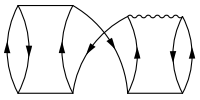
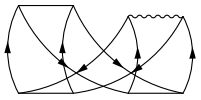
$$ E_{\mathrm{LPFD}} = E_{\mathrm{VCCD}} + \mathcal{O}(T^3). \tag{3.33} $$

In addition to this low-order correspondence, unlike CEPA(0), LPFD contains terms through all orders of the cluster amplitudes that are generated through the powers of the $\mathbf{\Delta}$ matrix in the binomial expansion above, and an examination of the $\mathcal{O}(T^3)$ terms yields additional insight. These odd-order terms are generated by the insertion of the expansion of the transformed amplitudes into the first

contribution to the LPFD correlation energy,

$$\begin{aligned}\langle \hat{H}\,{}_2\hat{T}\rangle &= \frac{1}{4}\langle ij||ab\rangle\, {}_2T_{ij}^{ab} \\ &= \frac{1}{8}\langle ij||ab\rangle(\mathbf{U}^{-1})_{ij}^{kl}T_{kl}^{ab} \\ &= \frac{1}{8}\langle ij||ab\rangle\sum_{n=0}^{\infty}\binom{-1}{n}(\mathbf{\Delta}^n)_{ij}^{kl}T_{kl}^{ab} \\ &\leftarrow -\frac{1}{8}\langle ij||ab\rangle\Delta_{ij}^{kl}T_{kl}^{ab}, \end{aligned} \tag{3.34}$$

where the left-arrow ($\leftarrow$) notation stands for "one contribution to the previous expression is". The other contribution, $\langle {}_1\hat{T}^\dagger(\hat{H}-\langle\hat{H}\rangle)\,{}_1\hat{T}\rangle$, generates the even-ordered terms. By inserting the definition of $\mathbf{\Delta}$,

$$\begin{aligned}\langle \hat{H}\,{}_2\hat{T}\rangle &\leftarrow -\frac{1}{8}\langle ij||ab\rangle\Delta_{ij}^{kl}T_{kl}^{ab} \\ &= -\frac{1}{8}T_{kl}^{ab}\langle ij||ab\rangle\left[\frac{1}{2}(1-\lambda)\left(\delta_i^k\eta_j^l-\delta_j^k\eta_i^l-\delta_i^l\eta_j^k+\delta_j^l\eta_i^k\right)+\lambda\eta_{ij}^{kl}\right] \\ &= -\frac{1}{16}(1-\lambda)T_{kl}^{ab}\langle ij||ab\rangle\left(\delta_i^k\eta_j^l-\delta_j^k\eta_i^l-\delta_i^l\eta_j^k+\delta_j^l\eta_i^k\right)-\frac{1}{8}\lambda T_{kl}^{ab}\langle ij||ab\rangle\eta_{ij}^{kl} \\ &= -\frac{1}{4}(1-\lambda)T_{ik}^{ab}\langle ij||ab\rangle\eta_j^k-\frac{1}{8}\lambda T_{kl}^{ab}\langle ij||ab\rangle\eta_{ij}^{kl} \\ &= \frac{1}{2}(1-\lambda)\ \text{  }\ -\lambda\ \text{  }\ , \end{aligned} \tag{3.35}$$

it becomes clear that the $\mathcal{O}(T^3)$ terms contributing to the LPFD correlation energy are exactly those terms, with the same λ-dependent weightings, that were established to be able to capture the $\mathcal{O}(T^3)$ contributions to VCCD in the 2-electron limit, due to the simplifications and mutual cancellations that occur between the VCCD terms in this limit. In particular, LPFD(0) generates the term $\frac{1}{2}\mathcal{B}$, and LPFD(+1) generates the term $-\mathcal{C}$. The best approximation to VCCD is clearly LPFD(-1), which generates $\mathcal{B}+\mathcal{C}$, exactly as the terms appear in VCCD itself. For reference, the algebraic interpretation of each of the $\mathcal{O}(T^3)$ VCCD terms is presented in Table 3.2.

Furthermore, it has already been observed that, through the binomial theorem, the transformed cluster amplitudes contain geometric series of terms that generate VCCD-like contributions to the energy through all orders of the cluster amplitudes. Since it has been demonstrated that LPFD is exact for 2 electrons, it must be true that the terms generated by LPFD, for different values of λ, form

representative subsets of the terms present in VCCD, and that, in particular, any omitted VCCD terms must vanish by mutual cancellation. The λ parameter controls the specific subset of VCCD terms generated, and the weightings of the terms within that set.

One final point, and one of fundamental importance for the rest of this thesis, is that, as was noted in a previous section, any finite truncation of $\langle e^{\hat{T}_2^\dagger} \hat{H} e^{\hat{T}} \rangle_L$ leads to a functional that cannot be exact for 2 electrons. This is because any truncation of this expression, which is infinite even in the limiting 2-electron case, renders it necessarily incomplete. LPFD overcomes this by constructing instead an infinite-order approximation that violates neither extensivity, nor an exact treatment of limiting systems. It therefore represents the first of a new family of quantum-chemical methods that has been named *Approximate Variational Coupled Cluster Theories*.

Chapter 4

The Linked Pair Functional II: Technical Details[†]

The Linked Pair Functional is a new and unique electronic structure method, possessing an extremely impressive array of methodological properties, and that has been shown to be deeply connected with the VCCD method. However, the novel part of the approach, the introduction of partial local normalization through matrix transformations of the cluster amplitudes, brings with it several complications that are only rarely encountered or discussed in the context of contemporary electronic structure models, such as the non-commutation of certain matrices that renders the correct minimization of the functional problematic, and concerns regarding the existence of the inverse and inverse square root of the transformation matrix, $\mathbf{U}$. It is the purpose of this chapter to discuss these technical details. In addition, the inclusion of the effects of single excitations into the LPFD scheme will be outlined.

[†]Relevant publications:

[94] J. B. Robinson and P. J. Knowles, J. Chem. Phys. **135**, 044113 (2011).

[95] J. B. Robinson and P. J. Knowles, J. Chem. Phys. **136**, 054114 (2012).

4.1 Minimization of the Functional

The ground-state energy in LPFD theory is the minimum of the functional with respect to the set of doubles-only cluster amplitudes, $\{T_{ab}^{ij}\}$, and a necessary condition for a minimum is that each of the partial first derivatives of the energy with respect to the set of cluster amplitudes vanish.

$$\frac{\partial E}{\partial T_{ab}^{ij}} = 0 \quad \forall \quad i,j,a,b \tag{4.1}$$

A computational implementation of LPFD requires an analytic expression for the above partial derivative. Using the Einstein summation convention[69], the differential of the energy may be written as follows.

$$dE = \frac{\partial E}{\partial T_{ab}^{ij}} dT_{ab}^{ij} \tag{4.2}$$

For technical reasons, the following definition is made,

$$\frac{1}{2} dE = \frac{1}{4} G_{ij}^{ab} dT_{ab}^{ij} \tag{4.3}$$

and the problem is thus to find G_{ij}^{ab}, commonly called "the residual". In this section, an expression for the residual of LPFD(+1) is derived. The restriction to LPFD(+1) is for simplicity of exposition, but the main result of this section, the differentiation of matrix powers, applies to all square symmetric matrices, a set to which all transformation matrices discussed in this thesis belong.

To begin, consider the differential of the LPFD energy, which may be written as follows,

$$dE = \frac{1}{2}\langle\Phi_0|\hat{H} d({}_2T_{ab}^{ij})|\Phi_{ij}^{ab}\rangle + \frac{1}{2}\langle\Phi_0|{}_1\hat{T}^\dagger(\hat{H} - \langle\hat{H}\rangle) d({}_1T_{ab}^{ij})|\Phi_{ij}^{ab}\rangle \tag{4.4}$$

$$= \frac{1}{2}\left({}_2\mathcal{V}_{ij}^{ab} d({}_2T_{ab}^{ij}) + {}_1\mathcal{V}_{ij}^{ab} d({}_1T_{ab}^{ij})\right) \tag{4.5}$$

$$= \frac{1}{2}\sum_{q=1}^{2} {}_q\mathcal{V}_{ij}^{ab} d({}_qT_{ab}^{ij}), \tag{4.6}$$

with the following definitions,

$${}_2\mathcal{V}_{ij}^{ab} = \langle\Phi_0|\hat{H}|\Phi_{ij}^{ab}\rangle \tag{4.7}$$

$${}_1\mathcal{V}_{ij}^{ab} = \langle\Phi_0|{}_1\hat{T}^\dagger(\hat{H} - \langle\hat{H}\rangle)|\Phi_{ij}^{ab}\rangle. \tag{4.8}$$

The transformed amplitudes differentiate as follows,

$$\begin{aligned} d({}_qT_{ab}^{ij}) &= \left(\mathbf{U}^{-\frac{q}{2}} d\mathbf{T} + d\left[\mathbf{U}^{-\frac{q}{2}}\right]\mathbf{T}\right)_{ab}^{ij} \\ &= \left(\mathbf{U}^{-\frac{q}{2}} d\mathbf{T}\right)_{ab}^{ij} + \left(d\left[\mathbf{U}^{-\frac{q}{2}}\right]\mathbf{T}\right)_{ab}^{ij}. \end{aligned} \tag{4.9}$$

The derivative of the LPFD(+1) transformation matrix, $\mathbf{U}$, is,

$$\begin{aligned} (d\mathbf{U})_{kl}^{ij} &= d(U_{kl}^{ij}) \\ &= \frac{1}{2} d\left(T_{cd}^{ij} T_{kl}^{cd}\right) \\ &= \frac{1}{2} dT_{cd}^{ij} T_{kl}^{cd} + \frac{1}{2} T_{cd}^{ij}\, dT_{kl}^{cd}, \end{aligned} \tag{4.10}$$

and a first attempt at differentiating its powers might be,

$$d\left(\mathbf{U}^x\right) = x\mathbf{U}^{x-1} d\mathbf{U} = x\, d\mathbf{U}\mathbf{U}^{x-1}. \tag{4.11}$$

The necessary working may then be carried out,

$$\begin{aligned} dE &= \frac{1}{2}\sum_{q=1}^{2} {}_q\mathcal{V}_{ij}^{ab} \left(\mathbf{U}^{-\frac{q}{2}}\, d\mathbf{T} + d\left[\mathbf{U}^{-\frac{q}{2}}\right]\mathbf{T}\right)_{ab}^{ij} \\ &= \frac{1}{4}\sum_{q=1}^{2} {}_q\mathcal{V}_{ij}^{ab} \left(\mathbf{U}^{-\frac{q}{2}}\right)_{kl}^{ij} dT_{ab}^{kl} - \frac{1}{8}\sum_{q=1}^{2} q\ {}_q\mathcal{V}_{ij}^{ab} dU_{kl}^{ij} \left(\mathbf{U}^{-\frac{q}{2}-1}\mathbf{T}\right)_{ab}^{kl} \\ &= \frac{1}{2}\sum_{q=1}^{2} {}_q\mathcal{A}_{ij}^{ab} dT_{ab}^{ij} - \frac{1}{8}\sum_{q=1}^{2} q\ {}_q\mathcal{V}_{ij}^{ab} dU_{kl}^{ij}\, {}_q\mathcal{B}_{ab}^{kl} \\ &= \frac{1}{2}\sum_{q=1}^{2} {}_q\mathcal{A}_{ij}^{ab} dT_{ab}^{ij} - \frac{1}{8}\sum_{q=1}^{2} q\left(dT_{ab}^{ij} T_{kl}^{ab} + T_{ab}^{ij} dT_{kl}^{ab}\right) {}_q\mathcal{C}_{ij}^{kl} \\ &= \frac{1}{2}\sum_{q=1}^{2} {}_q\mathcal{A}_{ij}^{ab} dT_{ab}^{ij} - \frac{1}{4}\sum_{q=1}^{2} q\ {}_q\mathcal{D}_{ij}^{kl} T_{kl}^{ab} dT_{ab}^{ij} \\ &= \frac{1}{2} dT_{ab}^{ij} \sum_{q=1}^{2} \left({}_q\mathcal{A}_{ij}^{ab} - q\ {}_q\mathcal{E}_{ij}^{ab}\right), \end{aligned}$$

and the residual straightforwardly read off to be,

$$G_{ij}^{ab} = \sum_{q=1}^{2} \left({}_q\mathcal{A}_{ij}^{ab} - q\ {}_q\mathcal{E}_{ij}^{ab}\right), \tag{4.12}$$

with the following definitions of the intermediate quantities,

$${}_q\mathcal{A}_{ij}^{ab} = \frac{1}{2}(\mathbf{U}^{-\frac{q}{2}})_{ij}^{kl}\, {}_q\mathcal{V}_{kl}^{ab} \tag{4.13}$$

$$ {}_q\mathcal{B}^{ij}_{ab} = \frac{1}{2}(\mathbf{U}^{-\frac{q}{2}-1})^{ij}_{kl} T^{kl}_{ab} \tag{4.14} $$

$$ {}_q\mathcal{C}^{ij}_{kl} = \frac{1}{2}\, {}_q\mathcal{V}^{ab}_{kl}\, {}_q\mathcal{B}^{ij}_{ab} \tag{4.15} $$

$$ {}_q\mathcal{D}^{ij}_{kl} = \frac{1}{2}\left({}_q\mathcal{C}^{ij}_{kl} + \left({}_q\mathcal{C}^{T}\right)^{ij}_{kl}\right) = \frac{1}{2}\left({}_q\mathcal{C}^{ij}_{kl} + {}_q\mathcal{C}^{kl}_{ij}\right) \tag{4.16} $$

$$ {}_q\mathcal{E}^{ab}_{ij} = \frac{1}{2}\, {}_q\mathcal{D}^{kl}_{ij} T^{ab}_{kl}\,. \tag{4.17} $$

Unfortunately, the above residual is *incorrect*; it yields a correct minimum for the LPFD(+1) functional *only* for a 2-electron system. The reason is the erroneous differentiation of matrix powers in Eq. 4.11; in general, there is no reason to expect a matrix, $\mathbf{U}$, to commute with its derivative, $d\mathbf{U}$. Take for example, the case $x = 2$.

$$ d\left(\mathbf{U}^2\right) = \mathbf{U}\, d\mathbf{U} + d\mathbf{U}\, \mathbf{U} \neq 2\mathbf{U}\, d\mathbf{U} \neq 2\, d\mathbf{U}\, \mathbf{U} \tag{4.18} $$

Thus, the above residual holds for the case of 2 electrons only because the matrix is then a scalar. In fact, a similar test for a 2-hole system, for which LPFD(+1) is also exact for reasons discussed in Chapter 6, makes the problem readily apparent, since then the LPFD(+1) matrix is *not* a scalar, but the method should still agree with CID. Using the above residual, it does not.

This section proves that the derivative of a real symmetric transformation matrix $\mathbf{U}$, a set to which all transformation matrices in this thesis belong, raised to some rational power x, can in fact be expressed in terms of the eigenvalues, $\{\epsilon_p\}$, and eigenvectors, $\{\mathbf{X}_p\}$, of the transformation matrix $\mathbf{U}$. The explicit analytic form of this result is given in Eq. 4.47. This result will then be used to derive the correct residual, $\{G^{ab}_{ij}\}$, for the case of LPFD(+1) in section 4.4.

Take $\mathbf{U}$ to be an $N_p \times N_p$ real symmetric matrix, where, for example, in the case of LPFD(+1), $N_p = \frac{N(N-1)}{2}$ is the number of electron pairs, and N is the number of electrons. The eigenvalues, $\{\epsilon_p\}$ of a real symmetric matrix are always real and its eigenvectors linearly independent, so that a set of orthonormal eigenvectors, $\{\mathbf{X}_p\}$ can always be constructed. The eigenproblem for $\mathbf{U}$ is,

$$ \mathbf{U}\mathbf{X}_p = \epsilon_p \mathbf{X}_p \quad p \in \{1, 2, \ldots, N_p\}, \tag{4.19} $$

or, more generally,

$$ \mathbf{U}\mathbf{X} = \mathbf{X}\epsilon, \tag{4.20} $$

where $\mathbf{X}$ is the matrix with the set of eigenvectors $\{\mathbf{X}_p\}$ as its columns and ϵ is the corresponding diagonal matrix of eigenvalues. The notation $\mathbf{X}_p$ thus denotes the p^{th} column of $\mathbf{X}$, and $\mathbf{X}_p^T$ is the transpose of this column, which is equivalent to the p^{th} row of $\mathbf{X}^T$. It also follows from the orthonormality of the eigenvectors that $\mathbf{X}$ is an orthogonal matrix,

$$\mathbf{X}\mathbf{X}^T = \mathbf{X}^T\mathbf{X} = \mathbf{1} \Leftrightarrow \mathbf{X}_p^T\mathbf{X}_q = \delta_{pq} \Leftrightarrow \mathbf{X}^t = \mathbf{X}^{-1}. \tag{4.21}$$

Since $\mathbf{U}$ is a real symmetric matrix it follows that $\mathbf{U}$ can always be diagonalized as follows,

$$\mathbf{U} = \mathbf{X}\epsilon\mathbf{X}^T. \tag{4.22}$$

It is noteworthy that the powers of $\mathbf{U}$ possess the same eigenvectors as $\mathbf{U}$ itself,

$$\mathbf{U}^x\mathbf{X}_p = \epsilon_p^x\mathbf{X}_p, \tag{4.23}$$

and that arbitrary rational powers of $\mathbf{U}$ may thus be defined in the standard way,

$$\mathbf{U}^x = \mathbf{X}\epsilon^x\mathbf{X}^T \quad x \in \mathbb{Q}. \tag{4.24}$$

In contrast to powers of $\mathbf{U}$, the powers of a diagonal matrix such as ϵ are very easy to calculate; the diagonal elements of a diagonal matrix raised to some power simply correspond to the powers of those diagonal elements. For example, cubing the 2×2 diagonal matrix with diagonal elements $\{3, 5\}$, produces another diagonal matrix with elements $\{27, 125\}$, which can be confirmed by repeated matrix multiplication. Therefore,

$$(\epsilon^x)_{pp} = (\epsilon_{pp})^x = \epsilon_p^x. \tag{4.25}$$

Eq. 4.24 allows an expression for the derivative of $\mathbf{U}^x$ to be found, as long as expressions for the derivatives of ϵ^x and $\mathbf{X}$ can first be found in terms of known quantities.

To begin, consider the problem of differentiating ϵ^x, the powers of the eigenvalue matrix. This is quite a trivial problem because ϵ is a diagonal matrix. As a consequence, $d\epsilon$ is also diagonal, since the off-diagonal elements of ϵ are fixed at zero and hence have no dependence on the amplitudes. Hence,

$$\begin{aligned} \epsilon_{pq} &= \epsilon_p\delta_{pq} \\ d\epsilon_{pq} &= d\epsilon_p\delta_{pq}. \end{aligned} \tag{4.26}$$

Now, since two square diagonal matrices (of the appropriate dimensions to permit matrix multiplication) always commute, $\epsilon^x \, d\epsilon = d\epsilon \, \epsilon^x$. It thus follows that

$$d\left(\epsilon^x\right) = x\epsilon^{x-1} \, d\epsilon = x \, d\epsilon \, \epsilon^{x-1}, \tag{4.27}$$

since commutation holds, and the eigenvalue matrix thus differentiates like a scalar. Furthermore, it is possible to derive an expression for $d\epsilon_p$. From projection of Eq. 4.19 onto $\mathbf{X}_p$,

$$\mathbf{X}_p^T \mathbf{U} \mathbf{X} = \epsilon_p \mathbf{X}_p^T \mathbf{X}_p = \epsilon_p \tag{4.28}$$

then performing the differentiation,

$$\begin{aligned} d\epsilon_p &= d\mathbf{X}_p^T \mathbf{U} \mathbf{X}_p + \mathbf{X}_p^T d\mathbf{U} \mathbf{X}_p + \mathbf{X}_p^T \mathbf{U} d\mathbf{X}_p \\ &= d\mathbf{X}_p^T \mathbf{U} \mathbf{X}_p + \mathbf{X}_p^T d\mathbf{U} \mathbf{X}_p + \left(\mathbf{U}\mathbf{X}_p\right)^T d\mathbf{X}_p \\ &= d\mathbf{X}^T \epsilon_p \mathbf{X}_p + \mathbf{X}_p^T d\mathbf{U} \mathbf{X}_p + \mathbf{X}_p^T \epsilon_p d\mathbf{X}_p \\ &= \mathbf{X}_p^T d\mathbf{U} \mathbf{X}_p + \epsilon_p \left(d\mathbf{X}_p^T \mathbf{X}_p + \mathbf{X}_p^T d\mathbf{X}_p\right). \end{aligned} \tag{4.29}$$

In fact, the second term can be shown not to contribute, since it is related to the differential of the orthonormality condition,

$$d\left(\mathbf{X}^T \mathbf{X}\right)_{pq} = d\mathbf{X}_p^T \mathbf{X}_q + \mathbf{X}_p^T \, d\mathbf{X}_q = d(\delta_{pq}) = 0, \tag{4.30}$$

which yields a very simple expression for $d\epsilon_p$,

$$d\epsilon_p = \mathbf{X}_p^T d\mathbf{U} \mathbf{X}_p, \tag{4.31}$$

which solves the problem of finding an expression for the differential of the eigenvalues. The elements of $d\mathbf{U}$ may be found by direct differentiation of the matrix, as was carried out above.

Next, an expression for the derivative of the eigenvector matrix, $\mathbf{X}$, must be found. This is a much more involved problem than the derivative of powers of the eigenvalue matrix. First, recall that the N_p eigenvectors of $\mathbf{U}$ are linearly independent. They therefore form a basis for $\mathbb{R}^{N_p}$. The columns of $d\mathbf{X}$ are vectors in $\mathbb{R}^{N_p}$, and thus each $d\mathbf{X}_p$ may be expanded in the basis of the eigenvectors of $\mathbf{U}$.

$$d\mathbf{X}_p = \sum_{q=1}^{N_p} \mathbf{X}_q \alpha_{qp} \tag{4.32}$$

This reduces the problem of finding $\{d\mathbf{X}_p\}$ to the problem of finding the scalar coefficients of the above expansion, $\{\alpha_{qp}\}$. Inserting this expansion into the differential of the orthonormality condition,

$$\begin{aligned} d\mathbf{X}_p^T\mathbf{X}_q + \mathbf{X}_p^T d\mathbf{X}_q &= \sum_{r=1}^{N_p} \left[\mathbf{X}_r^T \alpha_{rp}\mathbf{X}_q + \mathbf{X}_p^T\mathbf{X}_r\alpha_{rq}\right] \\ &= \sum_{r=1}^{N_p} \left[\alpha_{rp}\delta_{rq} + \alpha_{rq}\delta_{pr}\right] \\ &= \alpha_{qp} + \alpha_{pq} \\ &= 0, \end{aligned} \tag{4.33}$$

implies $\alpha_{pq} = -\alpha_{qp}$ so that α is a skew-symmetric matrix. In particular, $\alpha_{pp} = 0$. Hence, the expansion of $d\mathbf{X}_p$ may omit the diagonal term.

$$d\mathbf{X}_p = \sum_{\substack{q=1 \\ q\neq p}}^{N_p} \mathbf{X}_q \alpha_{qp} \tag{4.34}$$

Taking the differential of the eigenproblem,

$$d\mathbf{U}\mathbf{X}_p + \mathbf{U}d\mathbf{X}_p = d\epsilon_p\mathbf{X}_p + \epsilon_p\mathbf{X}_p, \tag{4.35}$$

rearranging,

$$\mathbf{U}d\mathbf{X}_p - \epsilon_p d\mathbf{X}_p = d\epsilon_p\mathbf{X}_p - d\mathbf{U}\mathbf{X}_p \tag{4.36}$$

and performing some further manipulation,

$$\begin{aligned} (\mathbf{U} - \epsilon_p\mathbf{1})\, d\mathbf{X}_p &= \mathbf{X}_p d\epsilon_p - d\mathbf{U}\mathbf{X}_p \\ &= \mathbf{X}_p\mathbf{X}_p^T d\mathbf{U}\mathbf{X}_p - d\mathbf{U}\mathbf{X}_p \\ &= -(\mathbf{1} - \mathbf{X}_p\mathbf{X}_p^T) d\mathbf{U}\mathbf{X}_p, \end{aligned} \tag{4.37}$$

then inserting the expansion of $d\mathbf{X}_p$ into the left hand side,

$$\sum_{\substack{q=1 \\ q\neq p}}^{N_p} (\mathbf{U} - \epsilon_p\mathbf{1})\, \mathbf{X}_q\alpha_{qp} = -(\mathbf{1} - \mathbf{X}_p\mathbf{X}_p^T) d\mathbf{U}\mathbf{X}_p \tag{4.38}$$

and applying again the eigenproblem yields the following result,

$$\sum_{\substack{q=1 \\ q\neq p}}^{N_p} (\epsilon_q - \epsilon_p)\, \mathbf{X}_q\alpha_{qp} = -\left(\mathbf{1} - \mathbf{X}_p\mathbf{X}_p^T\right) d\mathbf{U}\, \mathbf{X}_p. \tag{4.39}$$

The case $p \neq q$ but $\epsilon_p = \epsilon_q$, which can occur if $\mathbf{U}$ possesses an eigenvalue with algebraic multiplicity greater than one (repeated eigenvalues), also vanishes from the LHS,

$$\sum_{\substack{q=1 \\ q\neq p, \epsilon_q \neq \epsilon_p}}^{N_p} (\epsilon_q - \epsilon_p)\, \mathbf{X}_q \alpha_{qp} = -\left(\mathbf{1} - \mathbf{X}_p \mathbf{X}_p^T\right) d\mathbf{U}\, \mathbf{X}_p. \tag{4.40}$$

Projection onto $\mathbf{X}_r^T$ with $r \neq p$ and $\epsilon_r \neq \epsilon_p$,

$$\sum_{\substack{q=1 \\ q\neq p, \epsilon_q \neq \epsilon_p}}^{N_p} (\epsilon_q - \epsilon_p)\, \mathbf{X}_r^T \mathbf{X}_q \alpha_{qp} = -\mathbf{X}_r^T d\mathbf{U}\, \mathbf{X}_p + \mathbf{X}_r^T \mathbf{X}_p \mathbf{X}_p^T d\mathbf{U}\, \mathbf{X}_p, \tag{4.41}$$

application of the orthonormality condition,

$$\sum_{\substack{q=1 \\ q\neq p, \epsilon_q \neq \epsilon_p}}^{N_p} (\epsilon_q - \epsilon_p)\, \delta_{rq} \alpha_{qp} = -\mathbf{X}_r^T d\mathbf{U}\, \mathbf{X}_p + \delta_{rp} \mathbf{X}_p^T d\mathbf{U}\, \mathbf{X}_p, \tag{4.42}$$

and further simplification,

$$(\epsilon_r - \epsilon_p)\, \alpha_{rp} = -\mathbf{X}_r^T d\mathbf{U}\, \mathbf{X}_p, \tag{4.43}$$

allows determination of the coefficients.

$$\alpha_{rp} = \frac{\mathbf{X}_r^T\, d\mathbf{U}\, \mathbf{X}_p}{\epsilon_p - \epsilon_r} \quad r, p \in \{1, 2, \ldots, N_p\}, r \neq p, \epsilon_r \neq \epsilon_p \tag{4.44}$$

The case $\alpha_{pp} = 0$ is already known, and the only remaining case that is unknown is α_{rp} when $r \neq p$ but $\epsilon_r = \epsilon_p$. This cannot be determined by projection onto the above equation since the desired α_{rp} vanishes from the LHS. However, these values turn out to be unnecessary.

Finally, with explicit forms for $d\epsilon_p$ and $d\mathbf{X}_p$, consider differentiating Eq. 4.24,

$$d\left(\mathbf{U}^x\right) = d\mathbf{X} \epsilon^x \mathbf{X}^T + \mathbf{X} \epsilon^x\, d\mathbf{X}^T + x \mathbf{X} \epsilon^{x-1}\, d\epsilon \mathbf{X}^T. \tag{4.45}$$

Noting that $d\left(\mathbf{X}^T\right) \equiv \left(d\mathbf{X}\right)^T$, and expanding $d\mathbf{X}$ as above gives the following result,

$$\left(d\left(\mathbf{U}^x\right)\right)_{ij} = x \sum_{\substack{l=1 \\ m=1}}^{N_p} (d\mathbf{U})_{lm} \sum_{k=1}^{N_p} \epsilon_k^{x-1} X_{ik} X_{lk} X_{mk} X_{jk} + \sum_{\substack{k=1 \\ l=1}}^{N_p} \left(\epsilon_k^x - \epsilon_l^x\right) X_{il} X_{jk} \alpha_{lk}. \tag{4.46}$$

Excluding again the cases $k = l$ and $k \neq l$ but $\epsilon_k = \epsilon_l$ from the second summation on the grounds that they do not contribute, then inserting the remaining (and known) values of $\{\alpha_{lk}\}$ gives the following, final result, for the derivative of rational powers of the matrix $\mathbf{U}$,

$$d\left(\mathbf{U}^x\right)_{ij} = x \sum_{k=1}^{N_p} X_{ik} X_{jk} \epsilon_k^{x-1} \sum_{\substack{m=1 \\ n=1}}^{N_p} X_{mk}\, dU_{mn} X_{nk} + \sum_{\substack{k=1,l=1 \\ k\neq l, \epsilon_k \neq \epsilon_l}}^{N_p} \frac{\epsilon_k^x - \epsilon_l^x}{\epsilon_k - \epsilon_l} X_{il} X_{jk} \sum_{\substack{m=1 \\ n=1}}^{N_p} X_{ml}\, dU_{mn} X_{nk}. \tag{4.47}$$

It is possible to show that this expression generates some sensible expressions, such as $d(\mathbf{U}^2) = \mathbf{U}d\mathbf{U} + d\mathbf{U}\mathbf{U}$, and $d(\mathbf{U}^{-1}) = -\mathbf{U}^{-1}d\mathbf{U}\mathbf{U}^{-1}$. A computer program that requires the result of Eq. 4.47 needs only to solve the eigenproblem of Eq. 4.20 in order to make use of it. The values $\{dU_{mn}\}$ may be obtained by straightforward differentiation of the unpowered matrix, as was done above. It should be noted that the special case of a transformation matrix that possesses degenerate eigenvalues is automatically accounted for by the above analysis, and within Eq. 4.47. This result therefore applies to all square symmetric matrices, the set to which all transformation matrices in this thesis belong.

At this stage, a complete derivation of the correct LPFD(+1) residual may be given. However, this is presented in an appendix to this chapter, such that the ongoing discussion of the technical aspects of LPF theory is not disrupted.

4.2 Positivity Considerations

In the previous section, it was demonstrated that the use of rational powers of the transformation matrix, $\mathbf{U}$, in order to formulate the LPFD approach, in particular $\mathbf{U}^{-1}$ and $\mathbf{U}^{-\frac{1}{2}}$, leads to complications for the minimization of the functional arising from the differentiation of these powered matrices, but that this is both a tractable and soluble problem. However, the use of these matrix powers leads to another, more severe complication; if either $\mathbf{U}^{-1}$ or $\mathbf{U}^{-\frac{1}{2}}$ fail to exist then the LPFD energy becomes incalculable, rendering the entire method useless.

The matrix $\mathbf{U}^{-1}$ is not particularly problematic, because it only fails to exist when the matrix $\mathbf{U}$ is singular, that is, when its determinant, $|\mathbf{U}|$, equals zero, or, equivalently, when any of the eigenvalues of $\mathbf{U}$ equal zero since the determinant of a matrix is equal to the product of its eigenvalues. Still, avoiding this possibility entirely would be ideal. The existence of $\mathbf{U}^{-\frac{1}{2}}$, however, is a much stronger condition on the matrix $\mathbf{U}$. From Eq. 4.24, it is apparent that a real matrix $\mathbf{U}^{-\frac{1}{2}}$ fails to exist if any of the eigenvalues, $\{\epsilon_i\}$, of the matrix $\mathbf{U}$ are less than or equal to zero. This imposes the very strict condition that, for both $\mathbf{U}^{-1}$ and $\mathbf{U}^{-\frac{1}{2}}$ to exist, all eigenvalues of $\mathbf{U}$ must be strictly positive. In mathematical nomenclature, a matrix with all eigenvalues strictly positive is called *positive-definite*. A matrix with only positive or zero eigenvalues is called *positive-semidefinite*.

Since the transformation matrix is defined as the sum of an identity matrix and another matrix, $\mathbf{\Delta}$,

$$U_{kl}^{ij} = \delta_{kl}^{ij} + \Delta_{kl}^{ij}, \tag{4.48}$$

this reduces to the problem of ensuring that $\mathbf{\Delta}$ is positive-semidefinite, since by writing $\mathbf{\Delta}$ in spectral form,

$$\mathbf{\Delta} = \mathbf{X}\epsilon\mathbf{X}^T, \tag{4.49}$$

and from the following working,

$$\begin{aligned} \mathbf{U} &= \mathbf{1} + \mathbf{\Delta} \\ &= \mathbf{1} + \mathbf{X}\epsilon\mathbf{X}^T \\ &= \mathbf{X}\mathbf{X}^T + \mathbf{X}\epsilon\mathbf{X}^T \\ &= \mathbf{X}\left(\mathbf{1} + \epsilon\right)\mathbf{X}^T \end{aligned} \tag{4.50}$$

it can be seen that the eigenvalues of $\mathbf{U}$ will simply be the eigenvalues of $\mathbf{\Delta}$ incremented by 1. Thus, if $\mathbf{\Delta}$ is positive-semidefinite then $\mathbf{U}$ will be positive-definite since the eigenvalues will each be greater than or equal to 1. If $\mathbf{\Delta}$ is not positive-semidefinite, however, then there is still no guarantee that $\mathbf{U}$ will be positive-definite, since the eigenvalues of $\mathbf{\Delta}$ may take any arbitrarily large negative value. Thus, in order for LPFD calculations to be possible for general systems, the associated $\mathbf{\Delta}$ matrix must be at least positive-semidefinite for all possible values of the cluster amplitudes from which it is constructed.

Unfortunately, the potential positivity of the matrix for $\lambda = -1$ can be disproved by the existence of numerical counter-examples. For example, the LPFD(-1) **U** matrix is found to contain negative eigenvalues, causing the matrix powering procedure to fail, for difluorine at large interatomic separations with the STO-3G basis. This precludes the use of LPFD(-1), the optimal LPFD method for the approximation of VCCD, for general calculations.

In fact, the only LPFD method that employs a transformation matrix that is rigorously demonstrable to be positive-definite is LPFD(+1). This is because, for LPFD(+1), $\Delta^{ij}_{kl} = \eta^{ij}_{kl}$, and the $\mathbf{\Delta}$ matrix is Gramian; a matrix $\mathbf{G}$ is said to be Gramian if each of its elements correspond to inner products formed from a set of vectors[60], $\{\mathbf{x}_i\}$,

$$G_{ij} = \langle \mathbf{x}_i, \mathbf{x}_j \rangle. \tag{4.51}$$

It can be proved that a Gramian matrix is always positive semi-definite. The density matrices in Eqs. 3.23 and 3.24, can be seen to be Gramian, since their elements may be written in the required inner product form.

$$\eta^i_j = \langle \hat{T}^\dagger j i^\dagger \hat{T} \rangle \tag{4.52}$$

$$\eta^{ij}_{kl} = \langle \hat{T}^\dagger k l j^\dagger i^\dagger \hat{T} \rangle \tag{4.53}$$

Any value of λ except $\lambda = +1$, however, makes use of a linear combination of these density matrices, which is not positive-semidefinite in general. Thus, the constraint of matrix positivity forces the elimination of all LPFD methods except LPFD(+1).

There is however, still the possibility of reformulating the definition of the LPFD transformed amplitudes to make direct use of the other positive-semidefinite density matrix, η^i_j. This new theory, for which the form of the functional is unchanged,

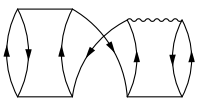
$$E = \langle \hat{H} \rangle + 2\langle \hat{H}\, {}_2\hat{T} \rangle + \langle {}_1\hat{T}^\dagger (\hat{H} - \langle \hat{H} \rangle)\, {}_1\hat{T} \rangle, \tag{4.54}$$

may be defined through a revised definition of the transformed amplitudes,

$$
\begin{aligned}
{}_q\hat{T}|\Phi_0\rangle &= \frac{1}{4}\,{}_qT^{ij}_{ab}|\Phi^{ab}_{ij}\rangle && (4.55)\\
{}_qT^{ij}_{ab} &= \frac{1}{2}(1-\tau_{ij})(\mathbf{U}^{-\frac{q}{2}})^i_k T^{kj}_{ab} && (4.56)\\
U^i_j &= \delta^i_j + \Delta^i_j && (4.57)\\
\Delta^i_j &= \eta^i_j = \frac{1}{2}T^{ik}_{ab}T^{ab}_{jk}, && (4.58)
\end{aligned}
$$

where the ground-state energy is again calculated as the minimum of the above energy functional with respect to the set of cluster amplitudes $\{T^{ij}_{ab}\}$. This is a method that is very much like performing LPFD(0) in a positive-definite way, since it generates the same $\mathcal{O}(T^3)$ term with the same weighting,

$$
\begin{aligned}
\frac{1}{4}\langle ij||ab\rangle\,{}_2T^{ab}_{ij} &= \frac{1}{4}\langle ij||ab\rangle\left[\frac{1}{2}(1-\tau_{ij})\left(\mathbf{U}^{-1}\right)^k_i T^{ab}_{kj}\right]\\
&= \frac{1}{8}\langle ij||ab\rangle(1-\tau_{ij})\sum_{n=0}^{\infty}\binom{-1}{n}(\eta^n)^k_i\,T^{ab}_{kj}\\
&\leftarrow -\frac{1}{8}\langle ij||ab\rangle(1-\tau_{ij})\eta^k_i T^{ab}_{kj}\\
&= -\frac{1}{4}\langle ij||ab\rangle\eta^k_i T^{ab}_{kj}\\
&= -\frac{1}{4}\langle ij||ab\rangle\eta^k_j T^{ab}_{ik}\\
&= +\frac{1}{2}
\end{aligned}
\qquad (4.59)
$$



and all of the noteworthy methodological properties of LPFD theory are preserved by this reformulation. In particular, it remains exact for 2 electrons because, although the transformation matrix is not a scalar in this limit, it is diagonal. For example, the following off-diagonal element,

$$
\begin{aligned}
U^e_{\bar{e}} &= \delta^e_{\bar{e}} + \frac{1}{2}T^{ek}_{ab}T^{ab}_{\bar{e}k}\\
&= \frac{1}{2}T^{ee}_{ab}T^{ab}_{\bar{e}e} + \frac{1}{2}T^{e\bar{e}}_{ab}T^{ab}_{\bar{e}\bar{e}}\\
&= 0, && (4.60)
\end{aligned}
$$

vanishes due to the fermionic antisymmetry of the cluster amplitudes. The diag-

onal elements contain the CID square norm,

$$\begin{aligned} U_e^e &= \delta_e^e + \frac{1}{2} T_{ab}^{ek} T_{ek}^{ab} \\ &= 1 + \frac{1}{2} T_{ab}^{ee} T_{ee}^{ab} + \frac{1}{2} T_{ab}^{e\bar{e}} T_{e\bar{e}}^{ab} \\ &= 1 + \frac{1}{2} T_{ab}^{e\bar{e}} T_{e\bar{e}}^{ab} \\ &= 1 + \langle \hat{T}_2^\dagger \hat{T}_2 \rangle, \end{aligned} \tag{4.61}$$

since for 2 electrons,

$$\begin{aligned} \langle \hat{T}_2^\dagger \hat{T}_2 \rangle &= \frac{1}{4} T_{ab}^{ij} T_{ij}^{ab} \\ &= \frac{1}{4} T_{ab}^{e\bar{e}} T_{e\bar{e}}^{ab} + \frac{1}{4} T_{ab}^{\bar{e}e} T_{\bar{e}e}^{ab} \\ &= \frac{1}{2} T_{ab}^{e\bar{e}} T_{e\bar{e}}^{ab}. \end{aligned} \tag{4.62}$$

Then, since the elements of the powers of a diagonal matrix correspond to the powers of the diagonal elements of that matrix, $(\epsilon^x)_i = (\epsilon_i)^x$, the transformed amplitudes correctly introduce the CID denominator,

$$\begin{aligned} {}_q T_{ab}^{e\bar{e}} &= \frac{1}{2} (\mathbf{U}^{-\frac{q}{2}})_k^e T_{ab}^{k\bar{e}} + \frac{1}{2} (\mathbf{U}^{-\frac{q}{2}})_k^{\bar{e}} T_{ab}^{ek} \\ &= \frac{1}{2} (\mathbf{U}^{-\frac{q}{2}})_e^e T_{ab}^{e\bar{e}} + \frac{1}{2} (\mathbf{U}^{-\frac{q}{2}})_{\bar{e}}^{\bar{e}} T_{ab}^{e\bar{e}} \\ &= \frac{1}{2} \frac{T_{ab}^{e\bar{e}}}{(1 + \langle \hat{T}_2^\dagger \hat{T}_2 \rangle)^{-\frac{q}{2}}} + \frac{1}{2} \frac{T_{ab}^{e\bar{e}}}{(1 + \langle \hat{T}_2^\dagger \hat{T}_2 \rangle)^{-\frac{q}{2}}} \\ &= \frac{T_{ab}^{e\bar{e}}}{(1 + \langle \hat{T}_2^\dagger \hat{T}_2 \rangle)^{-\frac{q}{2}}} \end{aligned} \tag{4.63}$$

Furthermore, this new formulation has the aesthetic property that $\mathbf{U}$ is now indexed by electrons, rather than by electron pairs, and so the matrix is significantly smaller in general; if N is the number of electrons then the number of unique electron pairs is $\frac{1}{2}N(N-1) \propto N^2$. Of course, this is *not* LPFD(0); the agreement between LPFD(0) and this new formulation of the theory is lost at high orders since the different definitions of matrix powers produce different terms. For example, in LPFD(0), $\eta_k^i \, \eta_l^j \, T_{ab}^{kl}$ is a contribution to $\frac{1}{2} \, (\mathbf{\Delta}^2)_{kl}^{ij} \, T_{ab}^{kl}$, but $(\eta^2)_j^i = \eta_k^i \, \eta_j^k$ and cannot produce this term. However, LPFD(0) and this new theory agree exactly to $\mathcal{O}(T^4)$, so, numerically, the differences are quite small. For this reason, and for convenience of description, this positive-definite approximate reformulation of LPFD(0) will itself be called LPFD(0), unless a specific distinction is made.

In conclusion, the positivity of the LPFD transformation matrix is only guaranteed in a few special cases, such as for the parameter value $\lambda = +1$ for the original LPFD transformation matrix, or for the reformulated LPFD(0)-like theory. In particular, LPFD(-1), the best LPFD approximation to VCCD, violates the positivity condition and is therefore not suitable for application to arbitrary systems because there is no guarantee that the LPFD(-1) energy is calculable. Since LPFD(+1) and the reformulated approximate LPFD(0) theory are the only LPFD methods that satisfy the positivity criterion, they are the only LPFD methods that will be given further consideration.

4.3 The Treatment of Single Excitations

The Linked Pair Functional (LPFD) is a doubles-only Variational Coupled Cluster approximation, in the sense that it accounts only for a subset of the terms present in VCCD. However, it is an approximation to VCCSD that is truly desired; although it is the double excitations that are the leading-order contributions to the correlation energy, the single excitations are also important, since, in particular, it is CISD that is the correct answer for a 2-electron system, not CID.

A natural extension of LPFD to a hypothetical LPFSD theory would be to incorporate the single excitation amplitudes into either of the positive-definite transformation matrices of LPFD(0) or LPFD(+1) such that, in the 2-electron limit, the CISD denominator is generated for all terms contributing to the functional, and then to employ either matrix to construct transformed single excitation operators, in addition to the transformed doubles, in order to construct that functional. Unfortunately, it appears difficult to reconcile this extension with the desire to preserve all of the advantageous LPFD methodological properties. For example, the LPFD(0) transformation matrix has the following elements,

$$U_j^i = \delta_j^i + \frac{1}{2} T_{ab}^{ik} T_{jk}^{ab}, \tag{4.64}$$

and the obvious extension is,

$$U_j^i = \delta_j^i + T_a^i T_j^a + \frac{1}{2} T_{ab}^{ik} T_{jk}^{ab}. \tag{4.65}$$

However, the singles contribution to the CISD denominator,

$$\langle \hat{T}_1^\dagger \hat{T}_1 \rangle = T_a^i T_i^a, \tag{4.66}$$

reduces to the sum of two distinct contributions in the two-electron limit,

$$\langle \hat{T}_1^\dagger \hat{T}_1 \rangle = T_a^e T_e^a + T_a^{\bar{e}} T_{\bar{e}}^a, \tag{4.67}$$

but only one of these two contributions enters U_j^i in this limit; $T_a^e T_e^a$ in the case of U_e^e and $T_a^{\bar{e}} T_{\bar{e}}^a$ in the case of $U_{\bar{e}}^{\bar{e}}$. Thus, this theory would not be exact for a limiting system containing only 2 electrons.

In contrast, the inclusion of singles explicitly into an LPFD(+1) transformation matrix is possible; consider the following potential form of the LPFD(+1) transformation matrix,

$$U_{kl}^{ij} = \delta_{kl}^{ij} + (1 - \tau_{ij})(1 - \tau_{kl}) T_a^i T_k^a \delta_l^j + \frac{1}{2} T_{ab}^{ij} T_{kl}^{ab}. \tag{4.68}$$

Noting that $\delta_{\bar{e}}^e$ and $\delta_e^{\bar{e}}$ vanish, $U_{e\bar{e}}^{e\bar{e}}$ contains the correct contributions such that $U_{e\bar{e}}^{e\bar{e}} = 1 + \langle \hat{T}_1^\dagger \hat{T}_1 \rangle + \langle \hat{T}_2^\dagger \hat{T}_2 \rangle$ for a 2-electron system. The contraction of this matrix with a double excitation amplitude is unchanged from LPFD(+1) theory, and is therefore straightforward, but contraction of this matrix with a single excitation amplitude is problematic because, in order to perform the contraction, two of the indices of the matrix must be eliminated in some way, such as in $U_{kj}^{ij} T_a^k$, where a sum over j is, of course, implied by the Einstein summation convention. This contains contributions such $\delta_k^i T_a^j T_j^a T_a^k = T_a^j T_j^a T_a^i = \langle \hat{T}_1^\dagger \hat{T}_1 \rangle T_a^i$, which clearly generates unlinked terms, such that the energy would not scale physically with the size of the system, and the theory therefore would not possess the property of rigorous extensivity.

For the above reasons, the explicit inclusion of single excitations into an LPF scheme currently remains an open problem. There are alternatives, however. The Brueckner Coupled Cluster method[96], restricted to double excitations (BCCD), for example, is a close relative of the CCSD method, but omits single excitations and instead replaces the equation to be solved for the singles with a Brueckner

condition[97, 98].

$$\langle\Phi_0|\hat{H}e^{\hat{T}_2}|\Phi_0\rangle = E_{\mathrm{BCCD}} \tag{4.69}$$

$$\langle\Phi_i^a|\hat{H}e^{\hat{T}_2}|\Phi_0\rangle = 0 \tag{4.70}$$

$$\langle\Phi_{ij}^{ab}|\hat{H}e^{\hat{T}_2}|\Phi_0\rangle = T_{ij}^{ab}E_{\mathrm{BCCD}} \tag{4.71}$$

This treats single excitations by absorbing their effects into the underlying orbitals, exploiting the equivalence between orbital rotations and single excitations implied by the Thouless Theorem[99], which states that any two single-determinantal wavefunctions, $|\Phi\rangle$ and $|\Phi'\rangle$, may be related by $|\Phi'\rangle = e^{\hat{T}_1}|\Phi\rangle$ for an appropriate choice of $\hat{T}_1$. An equivalent Brueckner condition,

$$\langle\Phi_i^a|\hat{H}(1 + {}_1\hat{T}_2)|\Phi_0\rangle = 0, \tag{4.72}$$

can be applied to the LPFD method, and the ground-state energy calculated as the minimum of the functional constrained to satisfy this relationship. This defines the Brueckner Linked Pair Function Doubles (BLPFD) method.

Alternatively, one can assert that the LPFD functional should be minimized directly with respect to both the doubles cluster amplitudes and the orbitals. This defines the Optimized-orbital Linked Pair Functional Doubles (OLPFD) method. The precise mechanics of performing this minimization are quite complicated, but have been discussed extensively in the context of coupled-pair functionals by Kollmar and Heßelmann[100]. The result is that the condition for the stationarity of the functional is that the partial derivatives with respect to the doubles amplitudes should vanish,

$$\frac{\partial E}{\partial T_{ab}^{ij}} = 0 \quad \forall \quad i, j, a, b, \tag{4.73}$$

and that the orbital gradient, f_{ia}, as defined in Ref. [100] should also vanish,

$$f_{ia} = 0 \quad \forall \quad i, a. \tag{4.74}$$

OLPFD has the advantage that it is fully variational, making the calculation of properties simpler, since solution of the linear equations that determine the lagrangian multipliers[101] is unnecessary. However, BLPFD is computationally

cheaper since the recalculation of integrals with three external indices in each iteration can be avoided[102]. While neither of these methods are as aesthetic as a hypothetical LPFSD theory, both BLPFD and OLPFD correctly account for the effects of single excitations such that they are exactly equivalent to CISD for a 2-electron system, and preserve all of the other attractive methodological properties of LPFD theory discussed in the previous chapter. However, the use of the Brueckner condition can sometimes lead to some surprising failures[103], and therefore OLPFD is the preferred approach.

4.4 Appendix: The Correct LPFD(+1) Residual

With the result of Section 4.1 for the differentiation of the transformation matrices raised to arbitrary rational powers, Eq. 4.47, it is possible to find a correct expression for the residual of an LPFD method, and, for completeness, this is presented here. For convenience, the simplest case of LPFD(+1) is again taken. Returning to the following point in the derivation,

$$\begin{aligned} dE &= \frac{1}{2}\sum_{q=1}^{2} {}_q\mathcal{V}_{ij}^{ab}\left(\mathbf{U}^{-\frac{q}{2}}\,d\mathbf{T} + d(\mathbf{U}^{-\frac{q}{2}})\mathbf{T}\right)_{ab}^{ij} \\ &= \frac{1}{2}\sum_{q=1}^{2} {}_q\mathcal{A}_{ij}^{ab} dT_{ab}^{ij} + \frac{1}{2}\sum_{q=1}^{2} {}_q\mathcal{B}_{ij}^{kl} d(\mathbf{U}^{-\frac{q}{2}})_{kl}^{ij}, \end{aligned} \tag{4.75}$$

the working for the first term is essentially complete, so developing the second term is now the goal. First, it is noteworthy that this second term can be converted from a spin-orbital notation in which the Einstein summation convention is taken to apply, to summations over the unique electron pairs only, taking ij and kl, for example, to be labels for the electron pairs (such that $\mathbf{U}$ effectively has two indices, not 4), and from there to an explicit matrix form in which the result of the previous subsection is more readily applicable. It is noteworthy that although ${}_q\mathcal{B}$ is not symmetric, the overall value of the contribution is independent

of the way in which the indices of ${}_q\mathcal{B}$ and $\mathbf{U}$ are combined.

$$\begin{aligned}\frac{1}{2}\sum_{q=1}^{2}{}_q\mathcal{B}_{ij}^{kl}\,d(\mathbf{U}^{-\frac{q}{2}})_{kl}^{ij} &= 2\sum_{q=1}^{2}\sum_{\substack{j<i\\l<k}}^{N}{}_q\mathcal{B}_{ij}^{kl}\,d(\mathbf{U}^{-\frac{q}{2}})_{kl}^{ij}\\ &= 2\sum_{q=1}^{2}\sum_{\substack{i=1\\j=1}}^{N_p}{}_q\mathcal{B}_{ij}\,d(\mathbf{U}^{-\frac{q}{2}})_{ji}\\ &= 2\sum_{q=1}^{2}\sum_{\substack{i=1\\j=1}}^{N_p}{}_q\mathcal{B}_{ji}\,d(\mathbf{U}^{-\frac{q}{2}})_{ij}\\ &= 2\sum_{q=1}^{2}\sum_{\substack{i=1\\j=1}}^{N_p}{}_q\mathcal{B}_{ji}\,d(\mathbf{U}^{-\frac{q}{2}})_{ji} \qquad (4.76)\end{aligned}$$

Substituting in the appropriate expression, the derivation proceeds as follows, where a number of intermediate quantities have been defined. These are summarised below in Eqs. 4.81-4.90.

$$\begin{aligned}\frac{1}{2}\sum_{q=1}^{2}{}_q\mathcal{B}_{ij}^{kl}\,d(\mathbf{U}^{-\frac{q}{2}})_{kl}^{ij} &= 2\sum_{q=1}^{2}\sum_{\substack{i=1\\j=1}}^{N_p}{}_q\mathcal{B}_{ij}\,d(\mathbf{U}^{-\frac{q}{2}})_{ji}\\ =2\sum_{q=1}^{2}\sum_{\substack{i=1\\j=1}}^{N_p}{}_q\mathcal{B}_{ij}\sum_{\substack{m=1\\n=1}}^{N_p}dU_{mn}&\left[\sum_{\substack{k=1,l=1\\k\neq l,\epsilon_k\neq\epsilon_l}}^{N_p}\frac{\epsilon_k^{-\frac{q}{2}}-\epsilon_l^{-\frac{q}{2}}}{\epsilon_k-\epsilon_l}X_{jl}X_{ik}X_{ml}X_{nk}\right.\\ &\left.-\frac{q}{2}\sum_{k=1}^{N_p}\epsilon_k^{-\frac{q}{2}-1}X_{jk}X_{ik}X_{mk}X_{nk}\right]\\ =2\sum_{q=1}^{2}\sum_{\substack{i=1\\m=1\\n=1}}^{N_p}dU_{mn}&\left[\sum_{\substack{k=1,l=1\\k\neq l,\epsilon_k\neq\epsilon_l}}^{N_p}\frac{\epsilon_k^{-\frac{q}{2}}-\epsilon_l^{-\frac{q}{2}}}{\epsilon_k-\epsilon_l}X_{ik}X_{ml}X_{nk}\,{}_q\mathcal{C}_{il}\right.\\ &\left.-\frac{q}{2}\sum_{k=1}^{N_p}\epsilon_k^{-\frac{q}{2}-1}X_{ik}X_{mk}X_{nk}\,{}_q\mathcal{C}_{ik}\right]\\ =2\sum_{q=1}^{2}\sum_{\substack{i=1\\m=1\\n=1}}^{N_p}dU_{mn}&\left[\sum_{\substack{k=1,l=1\\k\neq l,\epsilon_k\neq\epsilon_l}}^{N_p}\frac{\epsilon_k^{-\frac{q}{2}}-\epsilon_l^{-\frac{q}{2}}}{\epsilon_k-\epsilon_l}X_{ml}X_{nk}\,{}_q\mathcal{D}_{lk}\right.\\ &\left.-\frac{q}{2}\sum_{k=1}^{N_p}\epsilon_k^{-\frac{q}{2}-1}X_{mk}X_{nk}\,{}_q\mathcal{D}_{kk}\right]\\ =2\sum_{q=1}^{2}\sum_{\substack{i=1\\m=1\\n=1}}^{N_p}dU_{mn}&\left[\sum_{k=1}^{N_p}X_{nk}\,{}_q\mathcal{F}_{km}-\tfrac{q}{2}\,{}_q\mathcal{E}_{mn}\right]\end{aligned}$$

$$
\begin{aligned}
&= 2 \sum_{q=1}^{2} \sum_{\substack{m=1 \\ n=1}}^{N_p} dU_{mn} \left[{}_q\mathcal{H}_{mn} - \tfrac{q}{2} {}_q\mathcal{E}_{mn} \right] \\
&= 2 \sum_{q=1}^{2} \sum_{\substack{m=1 \\ n=1}}^{N_p} dU_{mn} {}_q\mathcal{I}_{mn}
\end{aligned} \tag{4.77}
$$

At this point, it is possible to convert back to a spin-orbital notation and to substitute the derivative of the unpowered matrix,

$$
\begin{aligned}
\frac{1}{2} \sum_{q=1}^{2} {}_q\mathcal{B}_{ij}^{kl} \, d(\mathbf{U}^{-\frac{q}{2}})_{kl}^{ij} &= 2 \sum_{q=1}^{2} \sum_{\substack{m=1 \\ n=1}}^{N_p} dU_{mn} {}_q\mathcal{I}_{mn} \\
&= 2 \sum_{q=1}^{2} \sum_{\substack{m=1 \\ n=1}}^{N_p} dU_{n}^{m} {}_q\mathcal{I}_{m}^{n} \\
&= \frac{1}{2} \sum_{q=1}^{2} dU_{kl}^{ij} {}_q\mathcal{I}_{ij}^{kl} \\
&= \frac{1}{4} \sum_{q=1}^{2} {}_q\mathcal{I}_{ij}^{kl} \left(T_{ab}^{ij} dT_{kl}^{ab} + dT_{ab}^{ij} T_{kl}^{ab} \right) \\
&= \frac{1}{4} \sum_{q=1}^{2} \left[{}_q\mathcal{I}_{ij}^{kl} + \left({}_q\mathcal{I}^{T} \right)_{ij}^{kl} \right] T_{kl}^{ab} dT_{ab}^{ij} \\
&= \frac{1}{4} \sum_{q=1}^{2} {}_q\mathcal{J}_{ij}^{kl} T_{kl}^{ab} dT_{ab}^{ij} \\
&= \frac{1}{2} \sum_{q=1}^{2} {}_q\mathcal{K}_{ij}^{ab} dT_{ab}^{ij}.
\end{aligned} \tag{4.78}
$$

Combining this with the first term from above, the differential of the energy is found to be,

$$
dE = \frac{1}{2} \sum_{q=1}^{2} \left[{}_q\mathcal{A}_{ij}^{ab} + {}_q\mathcal{K}_{ij}^{ab} \right] dT_{ab}^{ij}, \tag{4.79}
$$

from which the correct residual can be determined to be,

$$
G_{ij}^{ab} = \sum_{q=1}^{2} \left[{}_q\mathcal{A}_{ij}^{ab} + {}_q\mathcal{K}_{ij}^{ab} \right]. \tag{4.80}
$$

The definitions of the intermediate quantities used are as follows.

$$
{}_q\mathcal{A}_{ij}^{ab} = \frac{1}{2} (\mathbf{U}^{-\frac{q}{2}})_{ij}^{kl} \, {}_q\mathcal{V}_{kl}^{ab} \tag{4.81}
$$

$$
{}_q\mathcal{B}_{ij}^{kl} = \frac{1}{2} \, {}_q\mathcal{V}_{ij}^{ab} T_{ab}^{kl} \tag{4.82}
$$

$$ {}_q\mathcal{C}_{ij} = \sum_{k=1}^{N_p} {}_q\mathcal{B}_{ik} X_{kj} \tag{4.83} $$

$$ {}_q\mathcal{D}_{ij} = \sum_{k=1}^{N_p} {}_q\mathcal{C}_{ki} X_{kj} \tag{4.84} $$

$$ {}_q\mathcal{E}_{ij} = \sum_{k=1}^{N_p} \epsilon_k^{-\frac{q}{2}-1} X_{ik} X_{jk\,q}\mathcal{D}_{kk} \tag{4.85} $$

$$ {}_q\mathcal{F}_{ij} = \sum_{k=1,k\neq i,\epsilon_k\neq\epsilon_i}^{N_p} \frac{\epsilon_i^{-\frac{q}{2}} - \epsilon_k^{-\frac{q}{2}}}{\epsilon_i - \epsilon_k} X_{jk\,q}\mathcal{D}_{ki} \tag{4.86} $$

$$ {}_q\mathcal{H}_{ij} = \sum_{k=1}^{N_p} X_{jk\,q}\mathcal{F}_{ki} \tag{4.87} $$

$$ {}_q\mathcal{I}_{ij} = {}_q\mathcal{H}_{ij} - \frac{q}{2}\,{}_q\mathcal{E}_{ij} \tag{4.88} $$

$$ {}_q\mathcal{J}_{ij} = {}_q\mathcal{I}_{ij} + {}_q\mathcal{I}_{ji} \Leftrightarrow {}_q\mathcal{J} = {}_q\mathcal{I} + {}_q\mathcal{I}^T \tag{4.89} $$

$$ {}_q\mathcal{K}_{ij}^{ab} = \frac{1}{2}\,{}_q\mathcal{J}_{ij}^{kl} T_{kl}^{ab} \tag{4.90} $$

Chapter 5

Approximate Variational Coupled Cluster Theory[†]

In the previous two chapters, the fundamental theory of the Linked Pair Functional Doubles quantum-chemical method was outlined, and it was discussed how the effects of single excitations may be captured by variational minimization of the LPFD energy also with respect to the orbitals, defining the Optimized-orbital Linked Pair Functional Doubles method. Furthermore, it was found that only OLPFD(+1) and the approximate OLPFD(0) methods satisfy the constraint that the LPFD transformation matrix must be positive-definite if the LPFD energy is to be calculable. At this stage, it is necessary to benchmark the performance of the OLPFD(0) and OLPFD(+1) methods against CCSD, VCCSD and FCI. BCCD results are also given as a measure of the degree to which the different treatments of single excitations affect the results. It should be noted, however, that the Brueckner orbitals are not always close to the variationally optimal orbitals [100], and Brueckner Linked Pair Functional Doubles calculations have additionally been carried out as checks, although, for simplicity, these are not presented here. These calculations have been performed with the Molpro package of *ab initio* quantum chemistry programs[86, 87].

[†]Relevant publication:

[94] J. B. Robinson and P. J. Knowles, J. Chem. Phys. **135**, 044113 (2011).

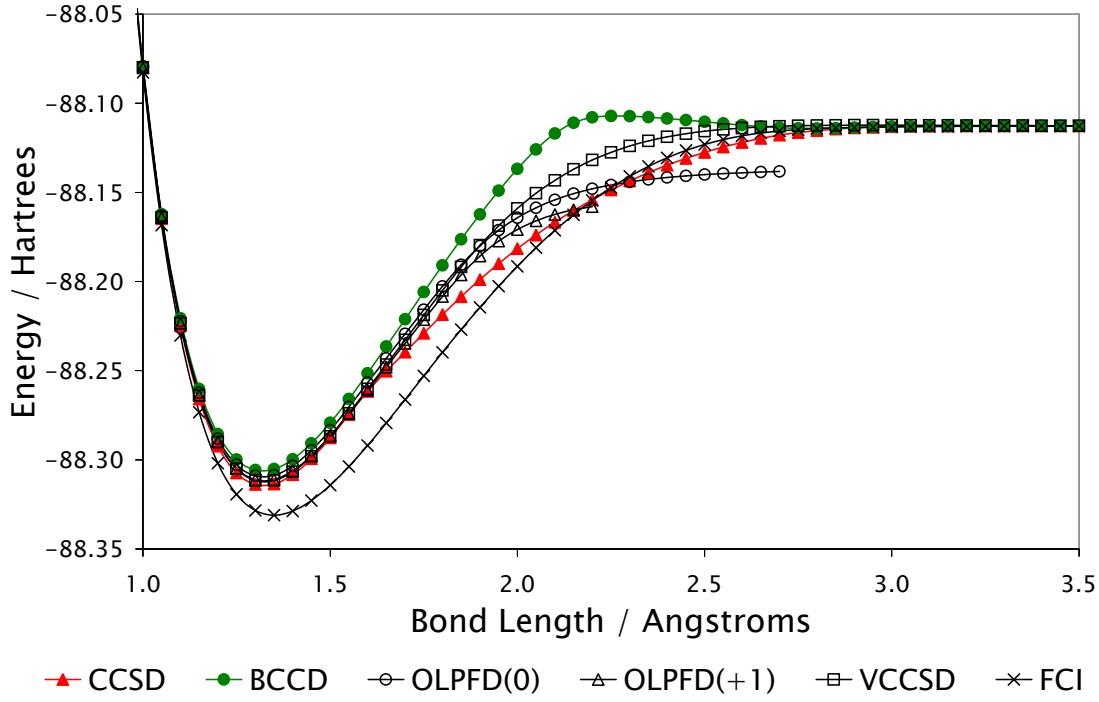


Figure 5.1: Calculated potential energy curves for BeO with the STO-3G basis set.

5.1 Preliminary OLPFD(0) and OLPFD(+1) Results

As a first example, consider the dissociation of the BeO molecule with the minimal STO-3G basis, for which results are given in Figure 5.1. The first noteworthy point is that neither of the TCC-based methods, CCSD and BCCD, perform particularly well; the CCSD curve is too shallow at long bond lengths, crossing the FCI curve at around 2.3Å, and the BCCD curve possesses an unphysical maximum. Neither curve therefore qualitatively mimics the shape of the FCI curve well. In contrast, the shape of the VCCSD curve is a significant improvement. Unfortunately, although both OLPFD(0) and OLPFD(+1) are closer to VCCSD at short bond lengths, they both eventually cross the FCI curve and in such a way that they appear to be approaching the wrong dissociation limits. By considering only this example, it clearly cannot be said that either of the positive-definite LPFD methods are better approximations of VCCSD than is CCSD, and clearly, this is not a good start.

Since these results are not particularly positive, consider next a model system in which four hydrogen atoms are arranged at the vertices of a rectangle that can be specified by two parameters. The parameter R, measured in angstroms, controls the distance of each H atom from the center of the rectangle. The parameter θ, measured in degrees, is the angle subtended at the center by lines to two neighbouring H nuclei. R therefore controls the overall size of the rectangle, and θ controls the distortion of the system from square ($\theta = 90°$), to rectangular geometry.

This system is of interest because the optimum Hartree-Fock reference determinant differs depending on whether $\theta < 90°$ or $\theta > 90°$. This can easily be understood because, if the hydrogen atoms are labelled A, B, C and D, then for $\theta \approx 0°$, atoms A and B, say, would be spatially distant from atoms C and D, and therefore the bonding would be A-B C-D, whereas for $\theta \approx 180°$, it would be, say, A and C nearby, and spatially distant from B and D, such that the bonding would instead be A-C B-D. The system therefore, by definition, possesses strong multireference character around the square geometry, since at least two determinants then become equally important to the description of the electronic structure. This is therefore a prototypical example of strong non-dynamic correlation, in which VCCSD is expected to be superior to CCSD. This system was first studied, at least in the context of VCCD benchmarking, by Van Voorhis and Head-Gordon[44], who found that around twice the 'equilibrium' value of R, the VCCD method is much more faithful to the shape of the FCI curve than CCSD. Analogous results, with R fixed and θ allowed to vary, and with the aug-cc-pVDZ basis set are supplied in Figure 5.2.

The first thing that should be apparent upon examination of this graph is that only the FCI curve is smooth through the square geometry. All of the single-reference systems display unphysical cusps at $\theta = 90°$, since, in these calculations, the reference determinant is allowed to vary. If the reference determinant is instead fixed, independent of θ, the single-reference methods become extremely poor and all predict asymmetric curves. The symmetry of this system about $\theta = 90°$ dictates that instead allowing the optimum reference determinant for a particular geometry to be used is the most faithful way to model this system with

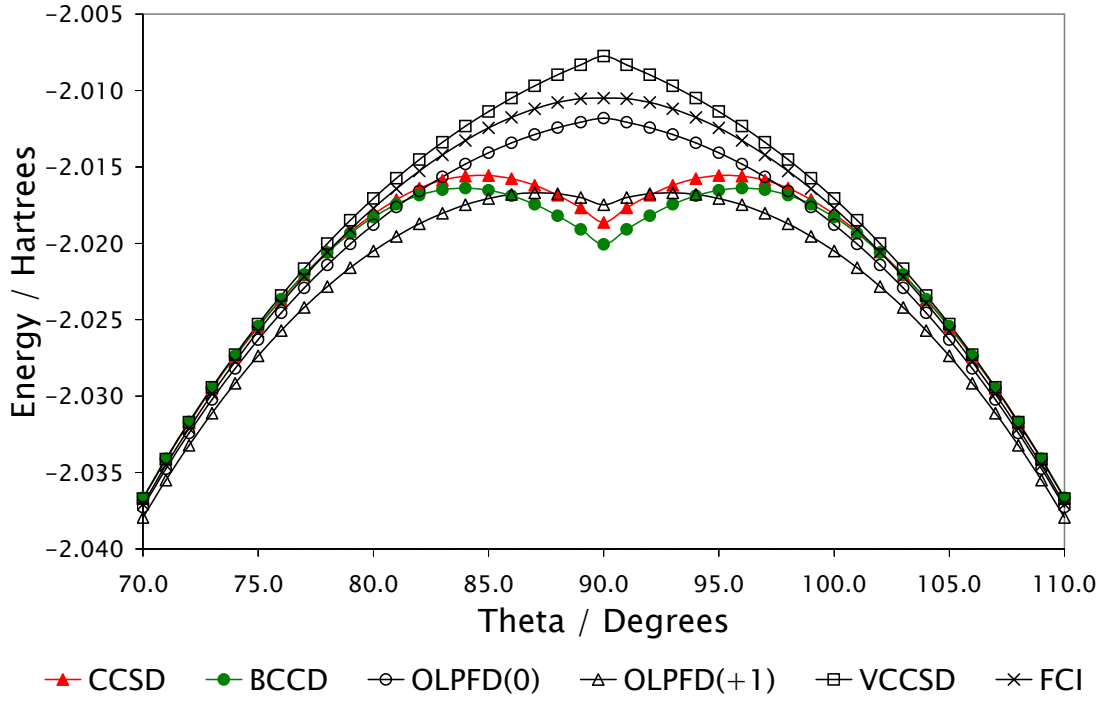


Figure 5.2: Calculated potential energy curves for H_4 with $R = 1.75$Å, and with the aug-cc-pVDZ basis set.

single-reference methods.

The distinction betweem the CCSD, BCCD and VCCSD methods in this system is significant; although the VCCSD curve is strongly cusped, the maximum at $\theta = 90°$ is analogous to the FCI curve, whereas both the CCSD and BCCD curves turn over around the square geometry, predicting less faithful, inverted cusps. The LPF methods show small improvements; the OLPFD(+1) curve is cusped in the same manner as the CCSD methods, but to a lesser degree, and the OLFPD(0) method is cusped in the same way as VCCSD, but is displaced below the FCI curve. It can therefore be said that the OLPFD approximation of VCCSD shows promise, but it is not strictly better than CCSD.

Finally, an examination of the dissociation of a multiply-bonded molecule is in order, since this is the type of system for which CCSD and related methods are known to perform poorly, and for which a more robust single-reference method would be highly desirable. The case of stretching the triple bond in acetylene (with the C-H bond length fixed at 1.06Å) is examined with the minimal STO-3G basis in Figure 5.3. A plot of the potential energy curve was previously given

in Figure 2.1, here errors relative to FCI are presented, with the effect of more clearly elucidating the behaviours of the different methods.

Even a cursory examination of Figure 5.3 reveals the huge difference between the behaviour of VCCSD, which remains above FCI throughout, and the other single-reference methods, which rapidly diverge to energies significantly below FCI for this system. It is particularly shocking, however, that while the TCC-based methods have reached an unphysical maximum and entered the non-variational turn-over phase by 1.7Å, the OLPFD methods, which have been constructed with the goal of better approximating VCCSD in situations such as this, in fact diverge more quickly, and at shorter bond lengths, with OLPFD(0), diverging from at least 1.6Å, performing only slightly better than OLPFD(+1), which diverges from at least 1.5Å.

In conclusion, while the OLPFD methods are extremely theoretically attractive, and despite their intimate relationship with the VCCSD functional, they do not solve the problem of constructing a single-reference method capable of correctly modelling strongly multireference phenomena, such as multiple bond breaking, and are not, in general, numerically close to VCCSD. This chapter rationalizes this problem, and proposes a potential solution such that a more accurate approximation to VCCSD can be constructed from the same LPFD mathematical and methodological principles.

5.2 A Corrected LPFD(0) Theory

The reason for the poor performance of the OLPFD methods, especially as approximations to VCCSD, can be explained through an examination of the low-order contributions to the correlation energy. The LPFD methods are exactly equivalent to VCCD to $\mathcal{O}(T^2)$ in the cluster amplitudes, or to third-order in Møller-Plesset perturbation theory, since this is the level of accuracy of the CEPA(0) functional that the LPFD amplitude transformations modify. At all higher orders, unlike CEPA(0), the LPFD methods contain subsets of the terms present in VCCD, and the use of these subsets is justified by the mutual cancel-

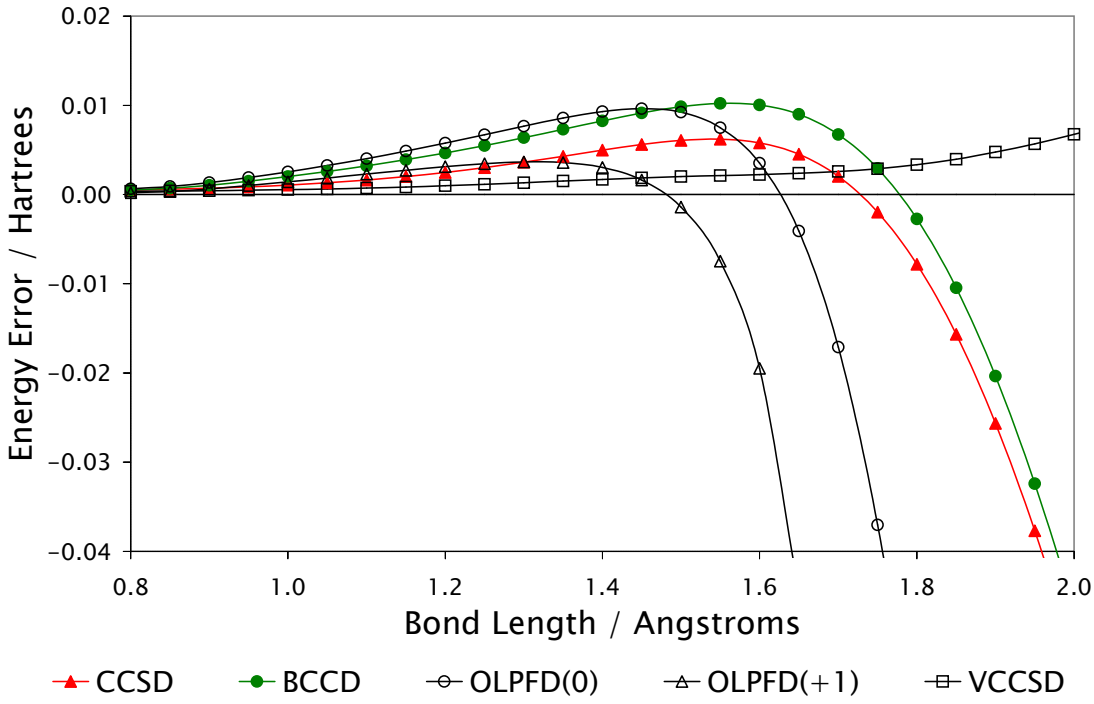


Figure 5.3: Calculated potential energy curves for the stretching of the carbon-carbon triple bond in acetylene, C_2H_2, with the C-H bond length fixed at 1.06Å, and with the STO-3G basis set.

lation of all other contributions in the 2-electron limit. For example, at $\mathcal{O}(T^3)$, the first order at which VCCD and LPFD differ, in the case of LPFD(0), only $\frac{1}{2}\mathcal{B}$ enters at this order, and the remaining omitted terms cancel for the case of 2 electrons,

$$\mathcal{A} + \frac{1}{2}\mathcal{B} + \mathcal{C} + \mathcal{D} = 0, \tag{5.1}$$

exactly as is implied by the relationships between the terms noted in Chapter 3. For the case of LPFD(+1), which generates diagram $\mathcal{C}$ with a factor of -1, the difference from VCCD at $\mathcal{O}(T^3)$ is,

$$\mathcal{A} + \mathcal{B} + 2\mathcal{C} + \mathcal{D} = 0, \tag{5.2}$$

which again vanishes for 2 electrons. These cancellations occur at all higher orders also.

However, the failure of the LPFD methods to achieve good numerical correspondence with VCCD can be understood as being due to the omission of so many low-order contributions, since the many omitted terms may become important for more general systems, containing more than 2 electrons. To put this in

perspective, LPFD(0) and LPFD(+1) both omit three of the four linked $\mathcal{O}(T^3)$ VCCD contributions, and recover the remaining term with an incorrect prefactor relative to how the term enters in VCCD. The correspondence with VCCD gets even worse at higher orders, and a scheme for correcting upon an LPFD starting point is required.

Since both the $\mathcal{B}$ and $\mathcal{C}$ terms are able to contribute in the 2-electron limit, they can be viewed as equally important, and a first correction to either LPFD(0) or LPFD(+1) should restore the balance of these two contributions, such that $\mathcal{B}+\mathcal{C}$ is obtained at $\mathcal{O}(T^3)$, exactly as the contributions appear in VCCD, and as in LPFD(-1), which was, of course, discarded on the grounds of positivity. Since the LPFD(0) method has been demonstrated to be slightly more accurate than LPFD(+1) in the numerical tests performed above, and since, as was discussed in Chapter 4, it is also slightly more simple to carry out computationally, the LPFD(0) method will be used as the underlying theory on which these low-order corrections are to be applied.

It is, of course, possible simply to add these low-order correction terms directly to the energy functional, but it is more satisfying to formulate their addition through the application of a new matrix transformation, since, for an appropriately chosen transformation, this ensures the preservation of all of the aesthetic LPFD methodological properties, and has the additional advantage that, if the transformation is applied multiplicatively, then contributions that couple the new transformation with the original $\mathbf{U}$ transformation will enter through all orders, and not just at low orders. The validity of these new infinite-order contributions can again be justified by the argument that they give a correct subset for 2-electrons, assuming that this can be demonstrated in a closed form.

Thus, the following transformation matrix is proposed,

$$ {}_qW_{kl}^{ij} = \delta_{kl}^{ij} + {}_qS_w\,\Omega_{kl}^{ij} \tag{5.3} $$

$$ \Omega_{kl}^{ij} = \eta_{kl}^{ij} - \frac{1}{2}(1-\tau_{ij})(1-\tau_{kl})\,\delta_k^i\,\eta_l^j. \tag{5.4} $$

This matrix reduces exactly to the identity, or rather a scalar equal to one, for

two electrons, since the $\mathbf{\Omega}$ matrix then vanishes,

$$\begin{aligned}
\Omega^{e\bar{e}}_{e\bar{e}} &= \eta^{e\bar{e}}_{e\bar{e}} - \frac{1}{2}\left(\delta^{e}_{e}\,\eta^{\bar{e}}_{\bar{e}} - \delta^{\bar{e}}_{e}\,\eta^{e}_{\bar{e}} - \delta^{e}_{\bar{e}}\,\eta^{\bar{e}}_{e} + \delta^{\bar{e}}_{\bar{e}}\,\eta^{e}_{e}\right) \\
&= \eta^{e\bar{e}}_{e\bar{e}} - \frac{1}{2}\left(\eta^{\bar{e}}_{\bar{e}} + \eta^{e}_{e}\right) \\
&= \langle \hat{T}_2^\dagger \hat{T}_2 \rangle - \frac{1}{2}\left(\langle \hat{T}_2^\dagger \hat{T}_2 \rangle + \langle \hat{T}_2^\dagger \hat{T}_2 \rangle\right) \\
&= 0. \qquad (5.5)
\end{aligned}$$

The transformation of the cluster amplitudes by this matrix may be defined as follows,

$$\left({}_{q}\mathbf{W}^{\,{}_{q}P_{w}}\,\mathbf{T}\right)^{ij}_{ab} = \frac{1}{2}\left({}_{q}\mathbf{W}^{\,{}_{q}P_{w}}\right)^{ij}_{kl} T^{kl}_{ab}, \qquad (5.6)$$

exactly the same as the $\mathbf{U}$ transformation.

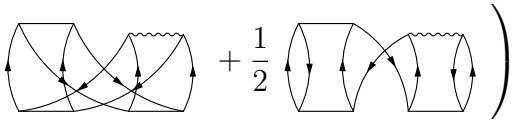
By expanding the powers of the $\mathbf{W}$ matrix through the binomial theorem, the leading-order contribution to the powered matrix can be seen to be the identity, such that the leading-order contribution to the $\mathbf{W}$-transformed amplitudes remains the untransformed cluster amplitudes, as is required for consistency. Applying this same procedure to examine the $\mathcal{O}(T^3)$ terms generated by this transformation gives the following,

$$\begin{aligned}
\left({}_{q}\mathbf{W}^{\,{}_{q}P_{w}}\,\mathbf{T}\right)^{ij}_{ab} &= \frac{1}{2}\left({}_{q}\mathbf{W}^{\,{}_{q}P_{w}}\right)^{ij}_{kl} T^{kl}_{ab} \\
&\leftarrow \frac{1}{2}\,{}_{q}P_{w}\,{}_{q}S_{w}\,\Omega^{ij}_{kl} T^{kl}_{ab} \\
&= \frac{1}{2}\,{}_{q}P_{w}\,{}_{q}S_{w}\left(\eta^{ij}_{kl} - \frac{1}{2}(1-\tau_{ij})(1-\tau_{kl})\,\delta^{i}_{k}\,\eta^{j}_{l}\right) T^{kl}_{ab} \\
&= \frac{1}{2}\,{}_{q}P_{w}\,{}_{q}S_{w}\left(\eta^{ij}_{kl} T^{kl}_{ab} - 2\eta^{j}_{l} T^{il}_{ab}\right) \\
&= {}_{q}P_{w}\,{}_{q}S_{w}\left(\frac{1}{2}\eta^{ij}_{kl} T^{kl}_{ab} - \eta^{j}_{k} T^{ik}_{ab}\right),
\end{aligned}$$

which, when inserted into the corresponding expression in the energy functional,

gives,

$$\begin{aligned}
\langle \hat{H}\, {}_2\hat{T} \rangle &= \frac{1}{4} \langle ij||ab \rangle \, {}_2T_{ij}^{ab} \\
\langle \hat{H}\, {}_2\hat{T} \rangle &= \frac{1}{4} \langle ij||ab \rangle \left({}_2\mathbf{W}^{\,{}_2P_w} \mathbf{T} \right)_{ij}^{ab} \\
&\leftarrow {}_2P_w \, {}_2S_w \langle ij||ab \rangle \left(\frac{1}{8} \eta_{ij}^{kl} T_{kl}^{ab} - \frac{1}{4} \eta_j^k T_{ik}^{ab} \right) \\
&= {}_2P_w \, {}_2S_w \left(\mathcal{C} + \frac{1}{2} \mathcal{B} \right) \\
&= {}_2P_w \, {}_2S_w \left(\quad + \frac{1}{2} \quad \right)
\end{aligned} \tag{5.7}$$



If the (at present unspecified) matrix powers, ${}_qP_w$, and matrix coefficients, ${}_qS_w$, satisfy,

$$ {}_2P_w \, {}_2S_w = +1, \tag{5.8}$$

or, more generally, for correct behavior for the $\mathcal{O}(T^4)$ 1-electron terms that will be necessary in Chapter 7,

$$ {}_qP_w \, {}_qS_w = +\frac{q}{2}, \tag{5.9}$$

the effect of this transformation is therefore to add $\frac{1}{2}\mathcal{B}+\mathcal{C}$ at $\mathcal{O}(T^3)$, complimenting the $\frac{1}{2}\mathcal{B}$ already present in LPFD(0) to give $\mathcal{B}+\mathcal{C}$, exactly as the terms enter VCCD.

Attention must now be given to the other two $\mathcal{O}(T^3)$ terms. One may be forgiven for thinking that since they do not contribute at all for a limiting system containing only 2 electrons, the $\mathcal{A}$ and $\mathcal{D}$ terms are 'less important' than either the $\mathcal{B}$ or $\mathcal{C}$ terms. This is false, however, because the cancellation of the terms is only exact for 2 electrons. In fact, in the general case, particularly when non-dynamic correlation is strong, the terms do not even approximately cancel, and the sum of the terms therefore represents an important contribution to the VCCD correlation energy. In addition, the $\mathcal{A}$ term is important in the 2-hole limit, which LPFD(0) does not treat correctly, and will be of interest in Chapter 6, and the $\mathcal{D}$ term is therefore at least as important, due to their 2-electron cancellation. Thus, these missing terms should be built into the prospective LPFD correction scheme.

Consider the following transformation matrix,

$$
\begin{aligned}
{}_qV^{ib}_{aj} &= \delta^{ib}_{aj} + {}_qS_v\,\Gamma^{ib}_{aj} && (5.10)\\
\Gamma^{ib}_{aj} &= 2\left(\delta^i_j\eta^b_a - \eta^{ib}_{aj}\right) && (5.11)\\
\eta^a_b &= \frac{1}{2}T^{ij}_{bc}\,T^{ac}_{ij} && (5.12)\\
\eta^{ib}_{aj} &= T^{ik}_{ac}\,T^{bc}_{jk}. && (5.13)
\end{aligned}
$$

This matrix again reduces to the identity for the case of 2 electrons since,

$$
\begin{aligned}
\Gamma^{eb}_{ae} &= 2\left(\delta^e_e\eta^b_a - \eta^{eb}_{ae}\right)\\
&= 2\left(\eta^b_a - \eta^{eb}_{ae}\right)\\
&= 2\left(\frac{1}{2}T^{bc}_{kl}T^{kl}_{ac} - T^{el}_{ac}T^{bc}_{el}\right)\\
&= 2\left(T^{bc}_{e\bar{e}}T^{e\bar{e}}_{ac} - T^{e\bar{e}}_{ac}T^{bc}_{e\bar{e}}\right)\\
&= 0, && (5.14)
\end{aligned}
$$

and,

$$
\begin{aligned}
\Gamma^{eb}_{a\bar{e}} &= 2\left(\delta^e_{\bar{e}}\eta^b_a - \eta^{eb}_{a\bar{e}}\right)\\
&= -2\eta^{eb}_{a\bar{e}}\\
&= -2T^{el}_{ac}T^{bc}_{\bar{e}l}\\
&= -2T^{ee}_{ac}T^{bc}_{\bar{e}e} - 2T^{e\bar{e}}_{ac}T^{bc}_{\bar{e}\bar{e}}\\
&= 0, && (5.15)
\end{aligned}
$$

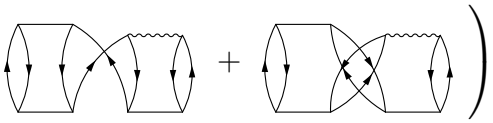
and when its powers are applied to transform the cluster amplitudes by the following scheme,

$$
\left({}_q\mathbf{V}^{\,qP_v}\,\mathbf{T}\right)^{ij}_{ab} = \frac{1}{4}(1-\tau_{ij})(1-\tau_{ab})\left({}_q\mathbf{V}^{\,qP_v}\right)^{ic}_{ak}\,T^{kj}_{cb} \qquad (5.16)
$$

which preserves the necessary fermionic antisymmetry, the following $\mathcal{O}(T^3)$ con-

tributions are generated.

$$\begin{aligned}
\frac{1}{4}\langle ij||ab\rangle\, {}_2T_{ij}^{ab} &= \frac{1}{4}\langle ij||ab\rangle \left({}_2\mathbf{V}^{\,2P_v}\mathbf{T}\right)_{ij}^{ab} \\
&\leftarrow \frac{1}{4}\,{}_2P_v\,{}_2S_v\langle ij||ab\rangle \left(\mathbf{\Gamma}\mathbf{T}\right)_{ij}^{ab} \\
&= \frac{1}{16}\,{}_2P_v\,{}_2S_v\langle ij||ab\rangle(1-\tau_{ij})(1-\tau_{ab})\Gamma_{ic}^{ak}T_{kj}^{cb} \\
&= \frac{1}{4}\,{}_2P_v\,{}_2S_v\langle ij||ab\rangle\Gamma_{ic}^{ak}T_{kj}^{cb} \\
&= \frac{1}{2}\,{}_2P_v\,{}_2S_v\langle ij||ab\rangle \left(\delta_i^k\eta_c^a - \eta_{ic}^{ak}\right) T_{kj}^{cb} \\
&= \frac{1}{2}\,{}_2P_v\,{}_2S_v\langle ij||ab\rangle\eta_c^a T_{ij}^{cb} - \frac{1}{2}\,{}_2P_v\,{}_2S_v\langle ij||ab\rangle\eta_{ic}^{ak}T_{kj}^{cb} \\
&= -\,{}_2P_v\,{}_2S_v(\mathcal{A}+\mathcal{D}) \\
&= -\,{}_2P_v\,{}_2S_v \left(\quad + \quad \right)
\end{aligned} \tag{5.17}$$



Therefore, if the (again, currently unset) matrix powers, ${}_qP_v$, and matrix coefficients, ${}_qS_v$, satisfy the following constraint,

$${}_qP_v\,{}_qS_v = -\frac{q}{2}, \tag{5.18}$$

then the correct VCCD terms, $\mathcal{A}+\mathcal{D}$, are generated at $\mathcal{O}(T^3)$, and also the correct 1-electron $\mathcal{O}(T^4)$ terms through $\langle {}_1\hat{T}^\dagger(\hat{H}-\langle\hat{H}\rangle)\,{}_1\hat{T}\rangle$. For completeness, an updated table (Table 5.1) of the $\mathcal{O}(T^3)$ VCCD terms is given, in which the contributions are written in terms of the newly-defined quantities.

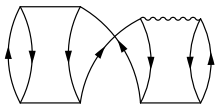
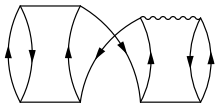
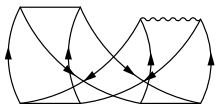
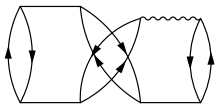
With these new transformations in hand, it is possible to fully define the corrected OLPFD(0) scheme, which will be named Approximate Variational Coupled Cluster Doubles (AVCCD) theory. The structure of the (previously LPFD) functional remains unchanged,

$$E_{\mathrm{AVCCD}} = \langle\hat{H}\rangle + 2\langle\hat{H}\,{}_2\hat{T}\rangle + \langle {}_1\hat{T}^\dagger(\hat{H}-\langle\hat{H}\rangle)\,{}_1\hat{T}\rangle, \tag{5.19}$$

and is to be minimized with respect to the untransformed cluster amplitudes in order to give a ground-state energy. The transformed cluster amplitudes, however, now take the following form,

$${}_qT_{ab}^{ij} = \left({}_q\mathbf{V}^{\,qP_v}\left({}_q\mathbf{W}^{\,qP_w}\left(\mathbf{U}^{-\frac{q}{2}}\mathbf{T}\right)\right)\right)_{ab}^{ij}. \tag{5.20}$$

Table 5.1: Linked $\mathcal{O}(T^3)$ contributions to VCCD, where $\langle \hat{T}^\dagger \hat{T} \rangle = \frac{1}{4} T_{ab}^{ij} T_{ij}^{ab}$, and where e and $\bar{e}$ label the two spinorbitals occupied in the reference wavefunction.

	$\frac{1}{2!}\langle \hat{T}_2^\dagger \hat{H} \hat{T}_2^2 \rangle_L$ Diagram	$\frac{1}{2!}\langle \hat{T}_2^\dagger \hat{H} \hat{T}_2^2 \rangle_L$ Term	2 Electrons
$\mathcal{A}$		$\frac{1}{4} T_{ij}^{ac}\, T_{cd}^{kl}\, T_{kl}^{db}\, \langle ij \vert\vert ab \rangle$ $= -\frac{1}{2} T_{ij}^{ac}\, \eta_c^b\, \langle ij \vert\vert ab \rangle$	$-T_{e\bar{e}}^{ac}\, \eta_c^b\, \langle e\bar{e} \vert\vert ab \rangle$
$\mathcal{B}$		$\frac{1}{4} T_{ik}^{ab}\, T_{cd}^{kl}\, T_{lj}^{cd}\, \langle ij \vert\vert ab \rangle$ $= -\frac{1}{2} T_{ik}^{ab}\, \eta_j^k\, \langle ij \vert\vert ab \rangle$	$-T_{e\bar{e}}^{ab} \langle \hat{T}^\dagger \hat{T} \rangle \langle e\bar{e} \vert\vert ab \rangle$
$\mathcal{C}$		$\frac{1}{16} T_{kl}^{ab}\, T_{cd}^{kl}\, T_{ij}^{cd}\, \langle ij \vert\vert ab \rangle$ $= \frac{1}{8} T_{kl}^{ab}\, \eta_{ij}^{kl}\, \langle ij \vert\vert ab \rangle$	$\frac{1}{2} T_{e\bar{e}}^{ab} \langle \hat{T}^\dagger \hat{T} \rangle \langle e\bar{e} \vert\vert ab \rangle$
$\mathcal{D}$		$\frac{1}{2} T_{ik}^{ac}\, T_{cd}^{kl}\, T_{lj}^{db}\, \langle ij \vert\vert ab \rangle$ $= \frac{1}{2} T_{ik}^{ac}\, \eta_{cj}^{kb}\, \langle ij \vert\vert ab \rangle$	$T_{e\bar{e}}^{ac}\, \eta_c^b\, \langle e\bar{e} \vert\vert ab \rangle$

As discussed above, the transformations are applied multiplicatively, or, more correctly, in a compound fashion, in which the transformations are applied successively. That is, first the LPFD(0) **U**-transformed amplitudes are constructed, then these amplitudes are themselves transformed by ${}_q\mathbf{W}$ to generate a new set of ${}_q\mathbf{W}$-**U**-transformed amplitudes, to which the ${}_q\mathbf{V}$ transformation is applied. The explicit form of each transformation has been outlined above. Since the new transformations are explicitly linked tensor quantities, still no unlinked terms are introduced, and the AVCCD method remains rigorously extensive, as well as invariant to rotations in the underlying orbital spaces. Since both of the corrective transformation matrices, ${}_q\mathbf{W}$ and ${}_q\mathbf{V}$, reduce to the identity matrix for the case of a 2-electron system, the AVCCD functional also reduces to the LPFD(0) functional, which is itself equal to the CID functional in this limit. AVCCD therefore remains exactly equivalent to CID for 2 electrons. Furthermore, the leading-order corrective terms enter linearly,

$$\begin{aligned}
&(\mathbf{1} + {}_qS_v\mathbf{\Gamma})^{\,{}_qP_v} (\mathbf{1} + {}_qS_w\mathbf{\Omega})^{\,{}_qP_w} (\mathbf{1} + \mathbf{\Delta})^{-\frac{q}{2}} \\
&= (\mathbf{1} + {}_qP_v\,{}_qS_v\mathbf{\Gamma} + \ldots)(\mathbf{1} + {}_qP_w\,{}_qS_w\mathbf{\Omega} + \ldots)(\mathbf{1} - \frac{q}{2}\mathbf{\Delta} + \ldots) \\
&= \mathbf{1} - \frac{q}{2}\mathbf{\Delta} + {}_qP_w\,{}_qS_w\mathbf{\Omega} + {}_qP_v\,{}_qS_v\mathbf{\Gamma} + \ldots,
\end{aligned} \tag{5.21}$$

and the AVCCD method therefore includes the $\mathcal{O}(T^3)$ terms exactly as they appear in VCCD.

When the AVCCD functional is also optimized with respect to the orbitals[100] (OAVCCD), or constrained by a Brueckner condition[96–98] (BAVCCD), in order to account for the effects of single excitations, the method possesses the following attractive theoretical properties.

- The ground-state energy is calculated by variational minimization of a functional. The error in a calculated energy is thus second-order in any remaining errors in the cluster amplitude parameters.
- The functional contains fully linked terms only, and is therefore rigorously extensive.
- It is exactly equivalent to FCI for a limiting 2-electron system.
- The energy is a scalar that is invariant to rotations in the underlying orbital spaces $\{\psi_i\}$ and $\{\psi_a\}$.
- The energy is not a true upper bound on the exact ground-state Schrödinger energy eigenvalue, but variational minimization of the functional gives a theory that satisfies the generalized Hellmann-Feynman theorem[67].
- The doubles-only theory is equivalent to VCCD to $\mathcal{O}(T^3)$, and, as a consequence, the one-electron $\mathcal{O}(T^4)$ terms are also constructed correctly.
- It is correct to third-order in Møller-Plesset perturbation theory[3] and omits only the terms containing triple excitations from fourth-order, the same as CCSD (see Chapter 7).
- The limiting computational complexity of the method is $\mathcal{O}(o^2v^4)$, the same as CCSD.

At this point, however, the powers and coefficients of the corrective transformation matrices remain unspecified. There are, in fact, no good theoretical reasons to justify any particular values, except that they must satisfy the constraints outlined above if the $\mathcal{O}(T^3)$ VCCD terms are to be generated with the correct weightings. AVCCD is therefore not a unique theory, a minor theoretical disadvantage. However, sensible values for the powers can be justified from the

constraint of positivity; neither of the corrective transformation matrices, ${}_q\mathbf{W}$ or ${}_q\mathbf{V}$, are strictly positive-definite in general. In particular, this prevents the use of $-\frac{q}{2}$ as the power of these matrices, analogous to the power of the $\mathbf{U}$ matrix, for reasons outlined in Chapter 4, and a restriction to positive integer powers must be made. Choosing powers of zero switches off the new transformations entirely, so, since it is the most computationally practical alternative, powers of $+1$ are proposed, since any other powers require additional matrix operations. Therefore,

$$ {}_qP_w = {}_qP_v = +1, \tag{5.22}$$

which, given the above constraints, allows the coefficients to be determined as,

$$ {}_qS_w = -{}_qS_v = +\frac{q}{2}. \tag{5.23}$$

With the AVCCD method fully specified, numerical benchmarking must be carried out, the topic of the next section.

5.3 Preliminary Results

The example systems of BeO, H_4 and acetylene, previously demonstrated a deficiency in the LPFD approximation of VCCD that has been attributed to the poor low-order correspondence of LPFD with VCCD. In particular, independent of the λ parameter, no LPFD method is able to match VCCD beyond $\mathcal{O}(T^2)$, since even LPFD(-1) omits the $\mathcal{A}$ and $\mathcal{D}$ terms at $\mathcal{O}(T^3)$. The AVCCD method, outlined in the previous section, however, by construction matches VCCD exactly to $\mathcal{O}(T^3)$, and to fourth-order in Møller-Plesset perturbation theory, by adding corrective terms to the LPFD(0) VCCD subset, without modification of the 2-electron behaviour. In this section, the above systems will be re-examined in order to determine to what level the numerical performance is improved by the improved approximation of VCCD.

The new BeO potential energy curves, for which the new AVCCD results, using orbital optimization to account for the effects of single excitations, have been added, are given in Figure 5.4 with the STO-3G basis set. At first glance,

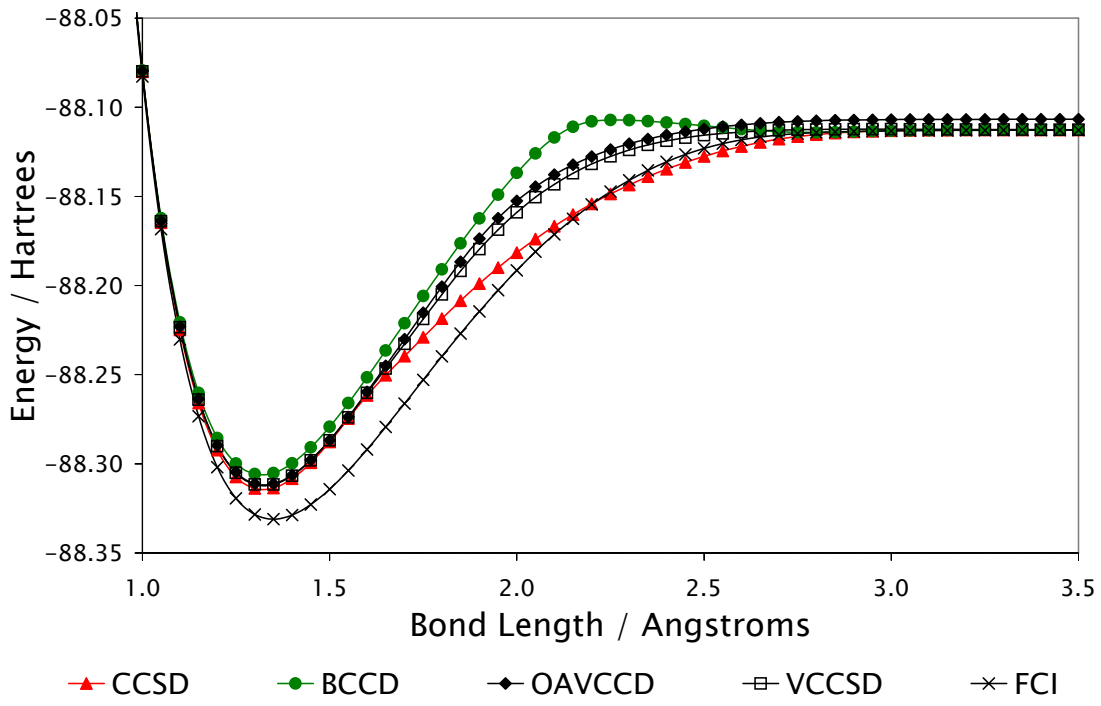


Figure 5.4: Calculated potential energy curves for BeO with the STO-3G basis set.

the AVCCD method appears to be performing better than either of the TCC methods; it neither crosses the FCI curve, as CCSD does, nor does it predict an unphysical maximum, as BCCD does. It also does not diverge from VCCSD at intermediate bond lengths, as was observed for the OLPFD methods in Figure 5.1, and appears to be a good approximation to VCCSD throughout. Unfortunately, however, unlike all of CCSD, BCCD and VCCSD, OAVCCD does not tend to the same limit, and does not become coincident with the FCI curve at long bond lengths, even when the bond appears to be fully broken. The AVCCD method is extensive, and exact for 2 electron systems, so this discrepancy cannot be explained by a shortcoming of its theoretical properties. Instead, it is likely to be associated with the omission of VCCD terms at higher orders, such that a further improvement to the approximation of VCCD may resolve the problem.

New results for the H_4 model system are given in Figure 5.5. Unlike the TCC potential energy curves, and the OLPFD potential energy curves given in Figure 5.2, the OAVCCD curve correctly mimics the VCCSD cusp shape. Furthermore, although it is too high in energy, the OAVCCD curve has the advantage that it remains above the FCI curve throughout, and the discrepancy from VCCSD

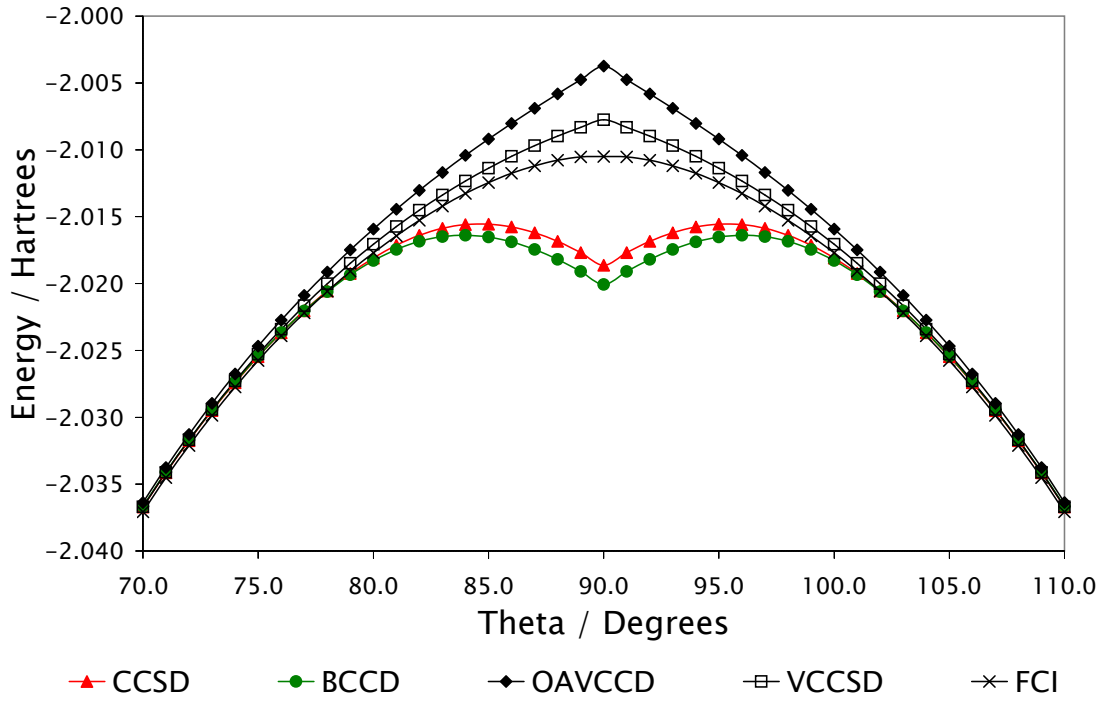


Figure 5.5: Calculated potential energy curves for H_4 with $R = 1.75$Å, and with the aug-cc-pVDZ basis set.

can again be attributed to the deficiency of the current VCCD approximation at orders higher than $\mathcal{O}(T^3)$.

Finally, the case of breaking the triple bond in acetylene is examined again in Figure 5.6. It was previously observed that the TCC methods predict an unphysical maximum in the potential energy curve for this system, followed by a non-variational breakdown to energies significantly below FCI, and that the LPFD methods perform even worse, as shown in Figure 5.3. The OAVCCD method, however, performs much better. In particular, it does not predict an unphysical maximum, and instead continues to increase monotonically at large bond lengths in a physically correct manner. Unfortunately, it is still significantly below the FCI curve, indicating that while it is potentially a more robust method than, for example, CCSD, it does not inherit sufficient upper bound character from its parent VCCD method to treat problems for which non-dynamic correlation is as strong as in this example.

In summary, these results indicate the AVCCD ansatz to be a remarkable improvement upon the LPFD methods for the treatment of systems for which non-

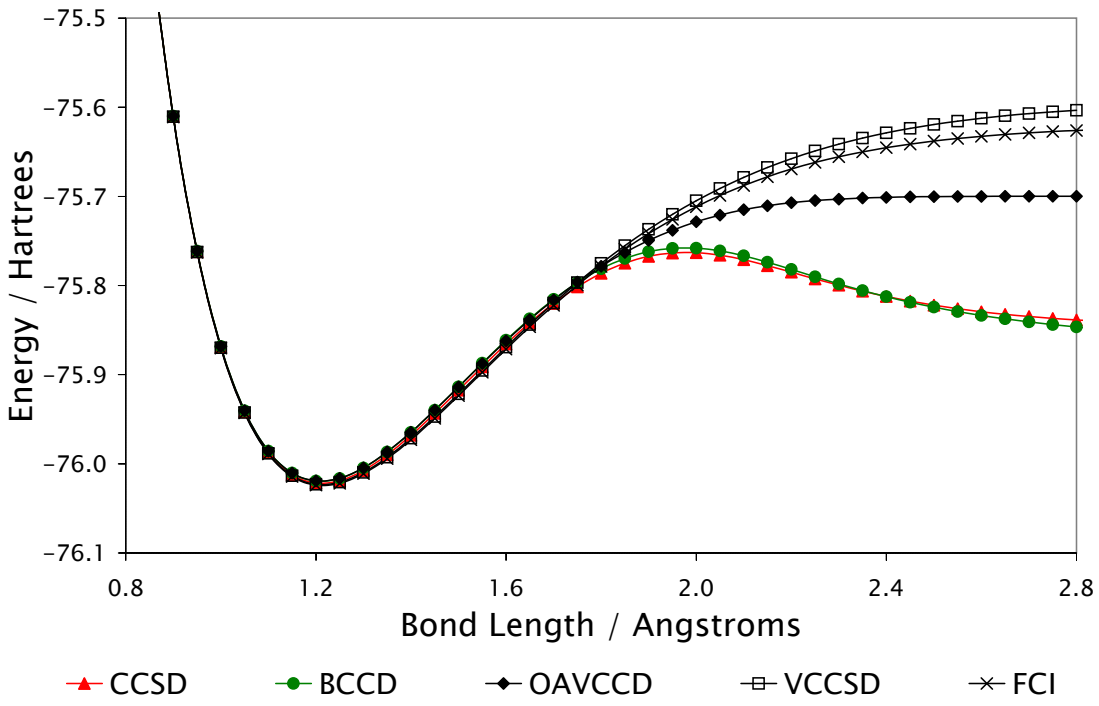


Figure 5.6: Calculated potential energy curves for the stretching of the carbon-carbon triple bond in acetylene, C_2H_2, with the C-H bond length fixed at 1.06Å, and with the STO-3G basis set.

dynamic correlation is strong, and the Hartree-Fock approximation breaks down, and shows promise as a method more suitable for the modelling of such systems than TCC. However, it does not fully resolve the problem of finding a robust and accurate single-reference method capable of correctly treating problems such as multiple bond breaking. If it is possible to further improve upon the AVCCD approximation to VCCD, this situation may be liable to change.

5.4 A Corrected LPFD(+1) Theory

In this chapter, it has already been discussed how low-order corrections to the LPFD(0) scheme can be constructed from the application of additional matrix transformations, such that the method agrees with VCCD to $\mathcal{O}(T^3)$. It has also been demonstrated that, although the resulting AVCCD method is more similar in spirit to VCCD than is CCD, and is more robust to the breakdown of the Hartree-Fock approximation that occurs in situations such as the dissociation of multiply-bonded molecules, it is still not robust enough to treat these problems

in a physically correct manner. Thus, a pertinent question is whether applying corrections to the other positive-definite LPFD method, LPFD(+1), performs any better. There is, in fact, a sound theoretical reason why this might be true; the LPFD(+1) method is exactly equivalent to CID, not just for 2 electrons, but for 2 holes (2 unoccupied virtual orbitals) also. This will be discussed further and rigorously proved in the following chapter, where preserving particle-hole symmetry will be promoted to a necessary criterion that an approximate VCC method must satisfy.

The correction scheme to be applied to LPFD(+1), in fact, requires only a trivial modification of the corrective transformations applied to LPFD(0). Consider the $\mathcal{O}(T^3)$ terms that must be added to LPFD(+1) to make it agree with VCCD exactly to this order; clearly, this is $\mathcal{A}+\mathcal{B}+2\mathcal{C}+\mathcal{D}$. With no modification, the ${}_q\mathbf{V}$ transformation can be used to account for $\mathcal{A}+\mathcal{D}$, and the $\mathbf{\Omega}$ matrix requires only to be multiplied by a factor of 2, such that the ${}_q\mathbf{W}$ transformation generates $\mathcal{B}+2\mathcal{C}$, instead of $\frac{1}{2}\mathcal{B}+\mathcal{C}$. However, if applied in the same multiplicative (or compound) fashion as the corrections to LPFD(0) above, although this hypothetical corrected LPFD(+1) theory would remain exact for 2 electrons, the corrective transformations break the hole-particle symmetry and prevent the reduction of the functional to CID for 2 holes. It is therefore suggested that the LPFD(+1)-based corrected transformed amplitudes are constructed as follows,

$$ {}_qT_{ab}^{ij} = (\mathbf{U}^{-\frac{q}{2}}\mathbf{T})_{ab}^{ij} - \frac{q}{2}(\mathbf{U}^{-\frac{q}{4}-\frac{1}{2}}(\mathbf{\Gamma}-\mathbf{\Omega})\mathbf{U}^{-\frac{q}{4}-\frac{1}{2}}\mathbf{T})_{ab}^{ij} \tag{5.24} $$

since, in this form, not only do the corrective transformation matrices vanish for 2 electrons, they also mutually cancel for 2 holes, such that LPFD(+1) (and therefore also CID) is recovered in both the 2-electron and 2-hole limits.

The proof of this cancellation is quite involved, and begins with an examination of the case of the application of the corrective transformations to the untransformed amplitudes, noting that the additional factor of 2 has been incor-

porated into the $\mathbf{\Omega}$ matrix,

$$\begin{aligned}((\mathbf{\Gamma}-\mathbf{\Omega})\,\mathbf{T})^{ij}_{ab} &= (\mathbf{\Gamma T}-\mathbf{\Omega T})^{ij}_{ab}\\ &= (\mathbf{\Gamma T})^{ij}_{ab}-(\mathbf{\Omega T})^{ij}_{ab}\\ &= \frac{1}{4}(1-\tau_{ij})(1-\tau_{ab})\Gamma^{ic}_{ak}T^{kj}_{cb}-\frac{1}{2}\Omega^{ij}_{kl}T^{kl}_{ab}\\ &= \frac{1}{4}(1-\tau_{ij})\Gamma^{ic}_{ak}T^{kj}_{cb}-\frac{1}{4}(1-\tau_{ij})\Gamma^{ic}_{bk}T^{kj}_{ca}-\frac{1}{2}\Omega^{ij}_{kl}T^{kl}_{ab}, \end{aligned} \tag{5.25}$$

and then to examine it specifically for 2 holes,

$$\begin{aligned}((\mathbf{\Gamma}-\mathbf{\Omega})\,\mathbf{T})^{ij}_{h\bar{h}} &= \frac{1}{4}(1-\tau_{ij})\Gamma^{ih}_{hk}T^{kj}_{h\bar{h}}+\frac{1}{4}(1-\tau_{ij})\Gamma^{i\bar{h}}_{\bar{h}k}T^{kj}_{h\bar{h}}-\frac{1}{2}\Omega^{ij}_{kl}T^{kl}_{h\bar{h}}\\ &= \frac{1}{4}(1-\tau_{ij})T^{kj}_{h\bar{h}}\left(\Gamma^{ih}_{hk}+\Gamma^{i\bar{h}}_{\bar{h}k}\right)-\frac{1}{2}\Omega^{ij}_{kl}T^{kl}_{h\bar{h}}\\ &= \frac{1}{2}(1-\tau_{ij})T^{kj}_{h\bar{h}}\left(\delta^i_k\eta^h_h+\delta^i_k\eta^{\bar{h}}_{\bar{h}}-\eta^{ih}_{hk}-\eta^{i\bar{h}}_{\bar{h}k}\right)\\ &\qquad -T^{kl}_{h\bar{h}}\left(\eta^{ij}_{kl}-\frac{1}{2}(1-\tau_{ij})(1-\tau_{kl})\delta^i_k\eta^j_l\right)\\ &= \frac{1}{2}(1-\tau_{ij})T^{ij}_{h\bar{h}}\left(\eta^h_h+\eta^{\bar{h}}_{\bar{h}}\right)-\frac{1}{2}(1-\tau_{ij})T^{kj}_{h\bar{h}}\left(\eta^{ih}_{hk}+\eta^{i\bar{h}}_{\bar{h}k}\right)\\ &\qquad -\eta^{ij}_{kl}T^{kl}_{h\bar{h}}+\frac{1}{2}T^{kl}_{h\bar{h}}(1-\tau_{ij})(1-\tau_{kl})\delta^i_k\eta^j_l. \end{aligned} \tag{5.26}$$

Then, because the following cancellations occur,

$$\begin{aligned}&\frac{1}{2}(1-\tau_{ij})T^{ij}_{h\bar{h}}\left(\eta^h_h+\eta^{\bar{h}}_{\bar{h}}\right)-\eta^{ij}_{kl}T^{kl}_{h\bar{h}}\\ &\qquad = \frac{1}{2}(1-\tau_{ij})T^{ij}_{h\bar{h}}T^{h\bar{h}}_{kl}T^{kl}_{h\bar{h}}-T^{ij}_{h\bar{h}}T^{h\bar{h}}_{kl}T^{kl}_{h\bar{h}}\\ &\qquad = T^{ij}_{h\bar{h}}T^{h\bar{h}}_{kl}T^{kl}_{h\bar{h}}-T^{ij}_{h\bar{h}}T^{h\bar{h}}_{kl}T^{kl}_{h\bar{h}}\\ &\qquad = 0 \end{aligned} \tag{5.27}$$

$$\begin{aligned}&\frac{1}{2}T^{kl}_{h\bar{h}}(1-\tau_{ij})(1-\tau_{kl})\delta^i_k\eta^j_l-\frac{1}{2}(1-\tau_{ij})T^{kj}_{h\bar{h}}\left(\eta^{ih}_{hk}+\eta^{i\bar{h}}_{\bar{h}k}\right)\\ &\qquad = (1-\tau_{ij})\eta^j_kT^{ik}_{h\bar{h}}-(1-\tau_{ij})T^{kj}_{h\bar{h}}T^{il}_{h\bar{h}}T^{h\bar{h}}_{kl}\\ &\qquad = (1-\tau_{ij})T^{kj}_{h\bar{h}}T^{h\bar{h}}_{kl}T^{il}_{h\bar{h}}-(1-\tau_{ij})T^{kj}_{h\bar{h}}T^{il}_{h\bar{h}}T^{h\bar{h}}_{kl}\\ &\qquad = 0, \end{aligned} \tag{5.28}$$

it is true that,

$$((\mathbf{\Gamma}-\mathbf{\Omega})\,\mathbf{T})^{ij}_{h\bar{h}}=0. \tag{5.29}$$

It is then additionally necessary to prove that the corrective transformations still cancel for 2 holes when applied to an amplitude related to the untransformed amplitudes by a non-negative integer power of $\mathbf{\Delta}$. This is trivially true since, in the 2-hole limit, $(\mathbf{\Delta}^x\mathbf{T})^{ij}_{h\bar{h}}$ is a scalar multiple of $T^{ij}_{h\bar{h}}$. Take $x = 1$, for example.

$$\begin{aligned}(\mathbf{\Delta T})^{ij}_{h\bar{h}} &= \frac{1}{2}T^{ij}_{h\bar{h}}T^{h\bar{h}}_{kl}T^{kl}_{h\bar{h}} \\ &= \kappa T^{ij}_{h\bar{h}} \quad \text{where} \quad \kappa = \frac{1}{2}T^{h\bar{h}}_{kl}T^{kl}_{h\bar{h}}\end{aligned} \tag{5.30}$$

Higher values of x follow trivially.

$$\begin{aligned}(\mathbf{\Delta}^x\mathbf{T})^{ij}_{h\bar{h}} &= \left(\mathbf{\Delta}^{x-1}\mathbf{\Delta T}\right)^{ij}_{h\bar{h}} \\ &= \frac{1}{2}\left(\mathbf{\Delta}^{x-1}\right)^{ij}_{mn}(\mathbf{\Delta T})^{mn}_{h\bar{h}} \\ &= \frac{1}{2}\kappa\left(\mathbf{\Delta}^{x-1}\right)^{ij}_{mn}T^{mn}_{h\bar{h}} \\ &= \kappa\left(\mathbf{\Delta}^{x-1}\mathbf{T}\right)^{ij}_{h\bar{h}} \\ &= \kappa^x T^{ij}_{h\bar{h}}\end{aligned} \tag{5.31}$$

Then the result follows easily.

$$\begin{aligned}((\mathbf{\Gamma}-\mathbf{\Omega})\mathbf{\Delta}^x T)^{ij}_{h\bar{h}} &= \kappa^x\left((\mathbf{\Gamma}-\mathbf{\Omega})\mathbf{T}\right)^{ij}_{h\bar{h}} \\ &= 0\end{aligned} \tag{5.32}$$

Finally, it possible to complete the proof that the corrective transformations cancel when they act on an amplitude related to the untransformed amplitudes by a rational power of $\mathbf{U}$. This is also fairly trivial, since any rational powers of $\mathbf{U}$ may be written as an infinite series of non-negative integer powers of $\mathbf{\Delta}$ through the binomial theorem. Each of the terms generated then vanishes in the 2-hole limit.

$$\begin{aligned}((\mathbf{\Gamma}-\mathbf{\Omega})\mathbf{U}^y\mathbf{T})^{ij}_{h\bar{h}} &= ((\mathbf{\Gamma}-\mathbf{\Omega})(1+\mathbf{\Delta})^y\mathbf{T})^{ij}_{h\bar{h}} \\ &= \sum_{x=0}^{\infty}\binom{y}{x}((\mathbf{\Gamma}-\mathbf{\Omega})\Delta^x\mathbf{T})^{ij}_{h\bar{h}} \qquad (5.33) \\ &= 0 \qquad (5.34)\end{aligned}$$

It is unfortunately true that, in practice, such a corrected LPFD(+1) scheme performs no better than the corrected LPFD(0) scheme for which numerical

benchmarking has been presented. There is, in fact, a very serious flaw in both of these schemes; adding only finite corrections to an infinite-order LPFD expression will be beneficial in general, but if the infinite-order expression becomes qualitatively wrong, as has been observed for the case of breaking the triple bond in acetylene, where both LPFD(0) and LPFD(+1) diverge unphysically, it becomes impossible to recover with only a finite-order correction. Decoupling a prospective Approximate Variational Coupled Cluster theory from the Linked Pair Functional in favour of a more robust infinite subset of VCCD terms will therefore be the subject of the following chapter.

Chapter 6

Quasi-Variational Coupled Cluster Theory[†]

It was discussed in the previous chapter how Approximate Variational Coupled Cluster theory yields, in principle, a good approximation, correct to $\mathcal{O}(T^3)$, to the VCCD method, and that this allows the OAVCCD method to predict a cusp shape for the rectangular H_4 model that is qualitatively similar to that of VCCSD. It is surprising, therefore, that OAVCCD performs only slightly better than CCSD in extreme examples such as breaking the triple bond in acetylene. VCCSD itself performs extremely well in such circumstances, which indicates that the OAVCCD method does not inherit sufficient upper bound character from VCCSD in order to treat these problems correctly; a superior approximation of VCCSD is required. The paradigm of the AVCC approach was to use either of the positive-definite LPF methods, LPFD(0) or LPFD(+1), to account for a representative subset of the terms present in VCCD through all orders of the cluster amplitudes such that the resulting method is exact for the appropriate limiting systems. Corrections were then added to further improve the approximation of VCCD at low orders. However, Fig. 6.1 shows the performances of both LPFD(0) and LPFD(+1) to be completely unsatisfactory for breaking the triple bond in acetylene. Thus, the failure of OAVCCD in this system can be traced back to the inadequacy of the

[†]Relevant publication:

[95] J. B. Robinson and P. J. Knowles, J. Chem. Phys. **136**, 054114 (2012).

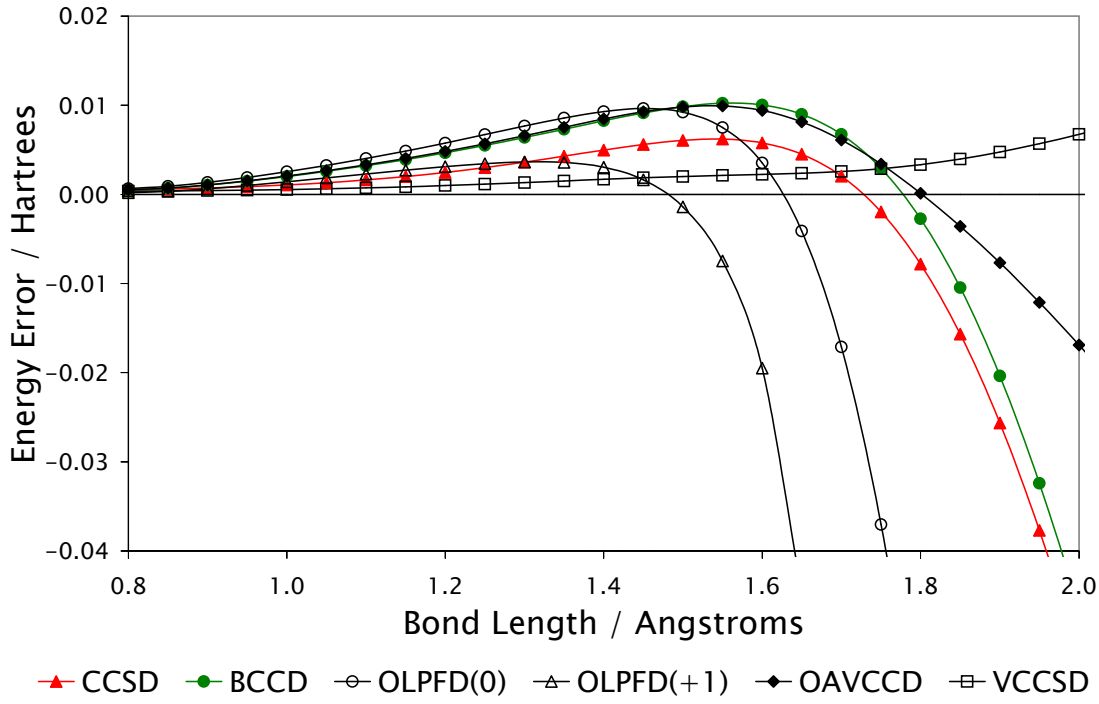


Figure 6.1: Errors relative to FCI in calculated energies for the stretching of the carbon-carbon triple bond in acetylene, C_2H_2, with the C-H bond length fixed at 1.06Å, and with the STO-3G basis set.

underlying approximation scheme of Linked Pair Functional theory.

This can be given a theoretical justification also; neither the LPFD(0) nor LPFD(+1) methods are particularly good approximations to VCCD from which to begin, each acquiring only one of the four unique $\mathcal{O}(T^3)$ terms, and, at higher orders, the situation can be expected to be even worse. More significantly, however, each of these methods generates only one *type* of VCCD term through all orders; LPFD(0) generates only terms that look like diagram $\mathcal{B}$ and LPFD(+1) generates only terms that look like diagram $\mathcal{C}$. This has been justified in the context of LPF theory by appealing to the internal mathematical structure of VCCD, that allows a series of, for example, $\mathcal{B}$-like terms, in the case of LPFD(0), or $\mathcal{C}$-like terms, in the case of LPFD(+1), to be correct for a system containing only 2 electrons due to the cancellation of the remaining terms. Fig. 6.2 shows what is meant by, for example, an $\mathcal{O}(T^4)$ $\mathcal{C}$-like term. However, for more general systems, the different terms present in VCCD can become equally important, and only by balancing these contributions correctly through all orders can a method truly approximate VCCD well. Put simply, if the underlying infinite-order ap-

proximation to VCCD becomes qualitatively wrong, as has been shown to occur in the case of acetylene, it becomes impossible to recover with only a finite-order correction, which dooms the AVCC approach to failure in such circumstances. For this reason, the approach of adding low-order corrections to an LPF method is fully discarded. It is important to point out, however, that the fundamental principle of modifying the CEPA(0) functional through matrix transformations of the cluster amplitudes such that partial local normalization is introduced, and accounting also for a subset of VCCD terms through all orders, remains valid. It is only the precise definition of the transformed cluster amplitudes that requires modification in order to achieve balanced contributions to VCCD through all orders.

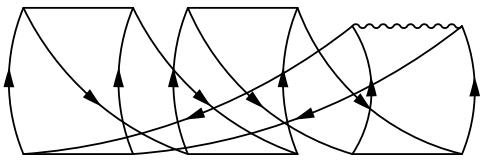


Figure 6.2: An example of an $\mathcal{O}(T^5)$ $\mathcal{C}$-like term that arises from $(\eta^2)^{kl}_{ij} T^{ab}_{kl} \langle ij||ab\rangle \propto \eta^{mn}_{ij} \eta^{kl}_{mn} T^{ab}_{kl} \langle ij||ab\rangle$

6.1 The QVCCD Method and its Transformed Amplitudes

The current goal is to construct a new definition for the transformed cluster amplitudes that, when inserted into the following functional,

$$E = \langle \hat{H} \rangle + 2\langle \hat{H}\, {}_2\hat{T} \rangle + \langle {}_1\hat{T}^\dagger (\hat{H} - \langle \hat{H} \rangle)\, {}_1\hat{T} \rangle, \tag{6.1}$$

not only yields a method that is identical with VCCD to $\mathcal{O}(T^3)$, but also balances higher-order terms such that problems involving strong static correlation may be treated by inheriting sufficient variational upper bound character from the parent VCCD method. This hypothetical theory is named *Quasi-Variational Coupled Cluster Doubles* (QVCCD).

In order to accomplish this goal, it would be desirable to construct a single positive-definite transformation matrix for each of the $\mathcal{O}(T^3)$ VCCD terms, such

that series of $\mathcal{A}$-like, $\mathcal{B}$-like, $\mathcal{C}$-like and $\mathcal{D}$-like terms may be generated separately and equivalently. In fact, it is possible to make use of the previously defined density matrices for this task,

$$ {}_{\mathcal{A}}\eta^{a}_{b} = \langle \hat{T}^{\dagger} a^{\dagger} b \hat{T} \rangle = \frac{1}{2} T^{ij}_{bc} \, T^{ac}_{ij} \tag{6.2} $$

$$ {}_{\mathcal{B}}\eta^{i}_{j} = \langle \hat{T}^{\dagger} j i^{\dagger} \hat{T} \rangle = \frac{1}{2} T^{ik}_{ab} \, T^{ab}_{jk} \tag{6.3} $$

$$ {}_{\mathcal{C}}\eta^{ij}_{kl} = \langle \hat{T}^{\dagger} k l j^{\dagger} i^{\dagger} \hat{T} \rangle = \frac{1}{2} T^{ij}_{ab} \, T^{ab}_{kl} \tag{6.4} $$

$$ {}_{\mathcal{D}}\eta^{ib}_{aj} = \langle \hat{T}^{\dagger} j b^{\dagger} a i^{\dagger} \hat{T} \rangle = T^{ik}_{ac} \, T^{bc}_{jk}, \tag{6.5} $$

where each density matrix has now been assigned a subscript denoting the $\mathcal{O}(T^3)$ VCCD term which it is responsible for generating. It should be noted that, since the elements of each density matrix can be written as inner products, then, when considered as true 2-index matrices with the following structures implied by the above inner products,

$$ {}_{\mathcal{A}}\eta_{a,b} = \langle \hat{T}^{\dagger} a^{\dagger} b \hat{T} \rangle \tag{6.6} $$

$$ {}_{\mathcal{B}}\eta_{i,j} = \langle \hat{T}^{\dagger} j i^{\dagger} \hat{T} \rangle \tag{6.7} $$

$$ {}_{\mathcal{C}}\eta_{ij,kl} = \langle \hat{T}^{\dagger} k l j^{\dagger} i^{\dagger} \hat{T} \rangle \quad ij, kl \in \{(m,n),\ m > n\} \tag{6.8} $$

$$ {}_{\mathcal{D}}\eta_{ia,jb} = \langle \hat{T}^{\dagger} j b^{\dagger} a i^{\dagger} \hat{T} \rangle \quad ia, jb \in \{(k,c)\}, \tag{6.9} $$

these density matrices are clearly all positive-semidefinite. The problem that the ${}_{q}\mathbf{W}$ and ${}_{q}\mathbf{V}$ transformation matrices are not positive-definite in general only arises because linear combinations of these density matrices are taken in order to construct transformation matrices. Thus, it is proposed that the transformed amplitudes take the form of a linear combination of several strictly positive-definite transformations, each constructed from a single one of the above density matrices, as is given below.

$$ \begin{aligned} {}_{q}T^{ij}_{ab} =& \alpha \left[\frac{1}{2} (1 - \tau_{ab}) \left({}_{\mathcal{A}}\mathbf{U}^{\mathcal{A}P} \right)^{c}_{a} T^{ij}_{cb} \right] \\ &+ \beta \left[\frac{1}{2} (1 - \tau_{ij}) \left({}_{\mathcal{B}}\mathbf{U}^{\mathcal{B}P} \right)^{i}_{k} T^{kj}_{ab} \right] \\ &+ \gamma \left[\frac{1}{2} \left({}_{\mathcal{C}}\mathbf{U}^{\mathcal{C}P} \right)^{ij}_{kl} T^{kl}_{ab} \right] \\ &+ \delta \left[\frac{1}{4} (1 - \tau_{ij})(1 - \tau_{ab}) \left({}_{\mathcal{D}}\mathbf{U}^{\mathcal{D}P} \right)^{ic}_{ak} T^{kj}_{cb} \right] \end{aligned} \tag{6.10} $$

Several currently unknown parameters have been introduced; α, β, γ and δ control the relative weightings of each of the transformations and ${}_{\mathcal{A}}P$, ${}_{\mathcal{B}}P$, ${}_{\mathcal{C}}P$ and ${}_{\mathcal{D}}P$ are the powers to which each of the transformation matrices is raised. The transformation matrices themselves are defined from the positive-semidefinite density matrices as follows,

$$ {}_{\mathcal{A}}U_b^a = \delta_b^a + \alpha' {}_{\mathcal{A}}\eta_b^a \tag{6.11} $$

$$ {}_{\mathcal{B}}U_j^i = \delta_j^i + \beta' {}_{\mathcal{B}}\eta_j^i \tag{6.12} $$

$$ {}_{\mathcal{C}}U_{kl}^{ij} = \delta_{kl}^{ij} + \gamma' {}_{\mathcal{C}}\eta_{kl}^{ij} \tag{6.13} $$

$$ {}_{\mathcal{D}}U_{aj}^{ib} = \delta_{aj}^{ib} + \delta' {}_{\mathcal{D}}\eta_{aj}^{ib} \tag{6.14} $$

where the parameters α', β', γ' and δ' act as prefactors on each of the density matrices. These transformation matrices are strictly positive-definite for any positive value of the primed prefactors. The purpose of introducing these parameters is to allow the transformed amplitudes sufficient flexibility such that the resulting theory can be made to meet the criteria of satisfying the important methodological properties that have been discussed previously. For example, it will be shown that several constraints on these parameters emerge from the criterion of an exact treatment of limiting systems containing only 2 electrons. Furthermore, it will be shown to be possible to tune the values of these parameters such that the method treats a system containing only 2 holes correctly, as CID and CCD already do, because, disregarding single excitations, the cluster operator $\hat{T} = \hat{T}_2$ is complete, not only for 2 electrons, but also in this limiting case. Finally, the third-order VCCD terms will be used to obtain further constraints from which unique values for each of these parameters can be derived.

Before proceeding to derive these important constraints, a preliminary constraint among the parameters α, β, γ and δ must first be established because, if the proposed functional is to be correct to $\mathcal{O}(T^3)$, it must necessarily be correct to $\mathcal{O}(T^2)$ first, and this means recovering CEPA(0) at low orders. Equivalently, this means that the leading-order contribution to the transformed amplitudes must be the untransformed amplitudes, since the QVCCD functional reduces to the CEPA(0) functional if the transformations are switched off. Expanding each of the transformation matrices through its binomial series, the leading-order contri-

butions to the new transformed amplitudes can be shown to be the untransformed amplitudes scaled by a multiplicative factor.

$$\begin{aligned}
{}_q\mathcal{T}_{ab}^{ij} =& \alpha \left[\frac{1}{2}(1-\tau_{ab})\left(\mathbf{1}+\mathcal{O}(T^2)\right)_a^c T_{cb}^{ij}\right] \\
&+\beta \left[\frac{1}{2}(1-\tau_{ij})\left(\mathbf{1}+\mathcal{O}(T^2)\right)_k^i T_{ab}^{kj}\right] \\
&+\gamma \left[\frac{1}{2}\left(\mathbf{1}+\mathcal{O}(T^2)\right)_{kl}^{ij} T_{ab}^{kl}\right] \\
&+\delta \left[\frac{1}{4}(1-\tau_{ij})(1-\tau_{ab})\left(\mathbf{1}+\mathcal{O}(T^2)\right)_{ak}^{ic} T_{cb}^{kj}\right] \\
=& \alpha \left[\frac{1}{2}(1-\tau_{ab})T_{ab}^{ij}\right] + \beta \left[\frac{1}{2}(1-\tau_{ij})T_{ab}^{ij}\right] \\
&+\gamma T_{ab}^{ij} + \delta \left[\frac{1}{4}(1-\tau_{ij})(1-\tau_{ab})T_{ab}^{ij}\right] + \mathcal{O}(T^3) \\
=& \left(\alpha+\beta+\gamma+\delta\right) T_{ab}^{ij} + \mathcal{O}(T^3)
\end{aligned} \tag{6.15}$$

This places the following constraint on the values of α, β, γ and δ.

$$\alpha+\beta+\gamma+\delta = 1 \tag{6.16}$$

6.2 Derivation I - Constraints Arising from the Limit of Two Electrons

In the limit of two electrons, some simplifying relationships emerge in the density matrices. Noting that in this limit, $\langle \hat{T}^\dagger \hat{T} \rangle \equiv \frac{1}{4} T_{ab}^{ij} T_{ij}^{ab} = \frac{1}{2} T_{ab}^{e\bar{e}} T_{e\bar{e}}^{ab}$, these are as follows.

$$\begin{aligned}
{}_{\mathcal{A}}\eta_b^a &= T_{bc}^{e\bar{e}} T_{e\bar{e}}^{ac} \\
{}_{\mathcal{B}}\eta_e^e &= \frac{1}{2} T_{ab}^{e\bar{e}} T_{e\bar{e}}^{ab} = {}_{\mathcal{B}}\eta_{\bar{e}}^{\bar{e}} = \langle \hat{T}^\dagger \hat{T} \rangle, \quad {}_{\mathcal{B}}\eta_{\bar{e}}^{e} = {}_{\mathcal{B}}\eta_e^{\bar{e}} = 0 \\
{}_{\mathcal{C}}\eta_{e\bar{e}}^{e\bar{e}} &= \frac{1}{2} T_{ab}^{e\bar{e}} T_{e\bar{e}}^{ab} = \langle \hat{T}^\dagger \hat{T} \rangle \\
{}_{\mathcal{D}}\eta_{be}^{ea} &= T_{bc}^{e\bar{e}} T_{e\bar{e}}^{ac} = {}_{\mathcal{D}}\eta_{b\bar{e}}^{\bar{e}a} = {}_{\mathcal{A}}\eta_b^a, \quad {}_{\mathcal{D}}\eta_{b\bar{e}}^{ea} = {}_{\mathcal{D}}\eta_{be}^{\bar{e}a} = 0
\end{aligned} \tag{6.17}$$

In this limit, ${}_{\mathcal{B}}\eta$ and ${}_{\mathcal{C}}\eta$ are both diagonal matrices (treating its indices as pair labels, ${}_{\mathcal{C}}\eta$ is, in fact, simply a scalar), and, as such, powers of the matrices corre-

spond simply to powers of the non-zero (diagonal) elements.

$$\begin{aligned}
({}_{\mathcal{B}}\eta^{x})^{i}_{j} &= \left({}_{\mathcal{B}}\eta^{i}_{j}\right)^{x}, \quad {}_{\mathcal{B}}\eta^{i}_{j} \neq 0 \\
({}_{\mathcal{C}}\eta^{x})^{ij}_{kl} &= \left({}_{\mathcal{C}}\eta^{ij}_{kl}\right)^{x}, \quad {}_{\mathcal{C}}\eta^{ij}_{kl} \neq 0
\end{aligned} \tag{6.18}$$

This behaviour persists when the corresponding identity contribution is added to form the transformation matrix. This has the consequence that the $\mathcal{B}$ and $\mathcal{C}$ terms both have the potential to introduce division by the CID norm in this limit.

$$\begin{aligned}
{}_{q}T^{e\bar{e}}_{ab} &\leftarrow \frac{\beta}{2}(1-\tau_{e\bar{e}})({}_{\mathcal{B}}\mathbf{U}^{\mathcal{B}P})^{e}_{k}T^{k\bar{e}}_{ab} + \frac{\gamma}{2}({}_{\mathcal{C}}\mathbf{U}^{\mathcal{C}P})^{e\bar{e}}_{kl}T^{kl}_{ab} \\
&= \frac{\beta}{2}(1-\tau_{e\bar{e}})({}_{\mathcal{B}}\mathbf{U}^{\mathcal{B}P})^{e}_{e}T^{e\bar{e}}_{ab} + \gamma({}_{\mathcal{C}}\mathbf{U}^{\mathcal{C}P})^{e\bar{e}}_{e\bar{e}}T^{e\bar{e}}_{ab} \\
&= \left(\beta(1+\beta'\langle\hat{T}^{\dagger}\hat{T}\rangle)^{\mathcal{B}P} + \gamma(1+\gamma'\langle\hat{T}^{\dagger}\hat{T}\rangle)^{\mathcal{C}P}\right)T^{e\bar{e}}_{ab}
\end{aligned} \tag{6.19}$$

If this is to happen correctly, such that the denominator introduced matches the CID norm exactly, then clearly it is required that the powers are,

$${}_{\mathcal{B}}P = {}_{\mathcal{C}}P = -\frac{q}{2} \tag{6.20}$$

such that ${}_{2}\hat{T}$ generates the norm with a power of -1, and that the product of the ${}_{1}\hat{T}^{\dagger}$ and ${}_{1}\hat{T}$ contributions, each of which should be the norm with a power of $-1/2$, also generates the norm with a power of -1 overall. Furthermore, the prefactor of $\langle\hat{T}^{\dagger}\hat{T}\rangle$ must be unity, or else the introduced denominator differs from that of CID,

$$\beta' = \gamma' = 1. \tag{6.21}$$

Finally, in order for the combined effect of the $\mathcal{B}$ and $\mathcal{C}$ terms to be correct, it is also required that the two contributions sum to one.

$$\beta + \gamma = 1 \tag{6.22}$$

Furthermore, it is useless for the CID denominator to be correctly generated through the $\mathcal{B}$ and $\mathcal{C}$ terms if the $\mathcal{A}$ and $\mathcal{D}$ terms cause erroneous contributions to enter. Fortunately, as can be seen from the simplification of the density matrices,

$${}_{\mathcal{D}}\eta^{ea}_{be} = {}_{\mathcal{D}}\eta^{\bar{e}a}_{b\bar{e}} = {}_{\mathcal{A}}\eta^{a}_{b} \tag{6.23}$$

there is an equality between the ${}_{\mathcal{D}}\eta$ and ${}_{\mathcal{A}}\eta$ matrices in this limit and thus the $\mathcal{A}$ and $\mathcal{D}$ terms can be arranged to cancel exactly. In order to do this, the above equivalence between the ${}_{\mathcal{A}}\eta$ and ${}_{\mathcal{D}}\eta$ matrices must first be shown to be true for integer powers of the matrices also.

$$\left({}_{\mathcal{D}}\eta^{n}\right)^{ea}_{be} = \left({}_{\mathcal{D}}\eta^{n}\right)^{\bar{e}a}_{b\bar{e}} = \left({}_{\mathcal{A}}\eta^{n}\right)^{a}_{b}, \quad n \in \mathbb{N}^{0} \tag{6.24}$$

This can be proved by induction; it has already been shown to be true for $n = 1$, and the case of $n = 0$ can also be trivially established.

$$\left({}_{\mathcal{D}}\eta^{0}\right)^{ea}_{be} = \delta^{ea}_{be} = \delta^{a}_{b} = \left({}_{\mathcal{A}}\eta^{0}\right)^{a}_{b} \tag{6.25}$$

For the inductive step, assume the relationship to hold for all integers up to n, then,

$$\begin{aligned}
\left({}_{\mathcal{D}}\eta^{n+1}\right)^{ea}_{be} &= \left({}_{\mathcal{D}}\eta\right)^{ec}_{bk}\left({}_{\mathcal{D}}\eta^{n}\right)^{ka}_{ce} \\
&= \left({}_{\mathcal{D}}\eta\right)^{ec}_{be}\left({}_{\mathcal{D}}\eta^{n}\right)^{ea}_{ce} \\
&= \left({}_{\mathcal{A}}\eta\right)^{c}_{b}\left({}_{\mathcal{A}}\eta^{n}\right)^{a}_{c} \\
&= \left({}_{\mathcal{A}}\eta^{n+1}\right)^{a}_{b}, \quad n \in \mathbb{N}^{0}
\end{aligned} \tag{6.26}$$

shows that it holds for $n + 1$, completing the proof. This result allows the constraints arising from cancellation of the $\mathcal{A}$ and $\mathcal{D}$ terms to be found.

$$\begin{aligned}
{}_{q}T^{e\bar{e}}_{ab} \leftarrow{}& \frac{1}{2}\alpha(1-\tau_{ab})\left({}_{\mathcal{A}}\mathbf{U}^{\mathcal{A}P}\right)^{c}_{a}T^{e\bar{e}}_{cb} + \frac{1}{4}\delta(1-\tau_{e\bar{e}})(1-\tau_{ab})\left({}_{\mathcal{D}}\mathbf{U}^{\mathcal{D}P}\right)^{ec}_{ae}T^{e\bar{e}}_{cb} \\
={}& \sum_{n=0}^{\infty}\binom{{}_{\mathcal{A}}P}{n}\left[\frac{1}{2}\alpha(1-\tau_{ab})\left(\alpha'^{n}\,{}_{\mathcal{A}}\eta^{n}\right)^{c}_{a}T^{e\bar{e}}_{cb}\right] \\
&+ \sum_{n=0}^{\infty}\binom{{}_{\mathcal{D}}P}{n}\left[\frac{1}{4}\delta(1-\tau_{e\bar{e}})(1-\tau_{ab})\left(\delta'^{n}\,{}_{\mathcal{D}}\eta^{n}\right)^{ec}_{ae}T^{e\bar{e}}_{cb}\right] \\
={}& \sum_{n=0}^{\infty}\binom{{}_{\mathcal{A}}P}{n}\left[\frac{1}{2}\alpha(1-\tau_{ab})\left(\alpha'^{n}\,{}_{\mathcal{A}}\eta^{n}\right)^{c}_{a}T^{e\bar{e}}_{cb}\right] \\
&+ \sum_{n=0}^{\infty}\binom{{}_{\mathcal{D}}P}{n}\left[\frac{1}{4}\delta(1-\tau_{e\bar{e}})(1-\tau_{ab})\left(\delta'^{n}\,{}_{\mathcal{A}}\eta^{n}\right)^{c}_{a}T^{e\bar{e}}_{cb}\right] \\
={}& \frac{1}{2}\sum_{n=0}^{\infty}\binom{{}_{\mathcal{A}}P}{n}\left[\alpha(1-\tau_{ab})\left(\alpha'^{n}\,{}_{\mathcal{A}}\eta^{n}\right)^{c}_{a}T^{e\bar{e}}_{cb}\right] \\
&+ \frac{1}{2}\sum_{n=0}^{\infty}\binom{{}_{\mathcal{D}}P}{n}\left[\delta(1-\tau_{ab})\left(\delta'^{n}\,{}_{\mathcal{A}}\eta^{n}\right)^{c}_{a}T^{e\bar{e}}_{cb}\right]
\end{aligned}$$

$$= \frac{1}{2}\alpha \sum_{n=0}^{\infty} \binom{{}_{\mathcal{A}}P}{n} \alpha'^{n} (1-\tau_{ab}) \left({}_{\mathcal{A}}\eta^{n}\right)_{a}^{c} T_{cb}^{e\bar{e}}$$

$$+ \frac{1}{2}\delta \sum_{n=0}^{\infty} \binom{{}_{\mathcal{D}}P}{n} \delta'^{n} (1-\tau_{ab}) \left({}_{\mathcal{A}}\eta^{n}\right)_{a}^{c} T_{cb}^{e\bar{e}}$$

$$= \frac{1}{2} \sum_{n=0}^{\infty} \left[\alpha\, \alpha'^{n} \binom{{}_{\mathcal{A}}P}{n} + \delta\, \delta'^{n} \binom{{}_{\mathcal{D}}P}{n} \right] (1-\tau_{ab}) \left({}_{\mathcal{A}}\eta^{n}\right)_{a}^{c} T_{cb}^{e\bar{e}} \tag{6.27}$$

Discarding trivial cases such as $\alpha = \delta = 0$ or $\alpha' = \delta' = 0$, which would switch off the $\mathcal{A}$ and $\mathcal{D}$ contributions permanently, in order for these series to cancel exactly, the following constraints must be applied,

$${}_{\mathcal{A}}P - {}_{\mathcal{D}}P = 0 \tag{6.28}$$

$$\alpha' - \delta' = 0 \tag{6.29}$$

$$\alpha + \delta = 0. \tag{6.30}$$

6.3 Derivation II - Constraints Arising from the Limit of Two Holes

Analogously to the limiting case of two electrons, one can examine the case of two holes, or two unoccupied spinorbitals in the reference wavefunction. This case occurs only in certain molecular examples and only when minimal basis sets are in use, but is nevertheless an important limiting case because, like the case of two electrons, a singles and doubles cluster operator is complete when only 2 holes exist, since at most two electrons can then be excited into the two virtual spinorbitals without violating the exclusion principle. Crucially, while CCD and VCCD are therefore equivalent to CID in both the limits of 2 electrons and 2 holes, LPFD(0) and AVCCD are correct only for 2 electrons, and not 2 holes. It would be especially theoretically attractive if correct behaviour in this limit could be additionally built into the QVCCD method. In order to do this, however, it becomes important to understand an additional property of LPFD(+1) and its associated transformation matrix, ${}_{\mathcal{C}}\mathbf{U}$, which is itself used in QVCCD theory.

First, some insight can be gained by considering once again the $\mathcal{O}(T^3)$ VCCD terms. Like the case of 2 electrons, simplifying relationships between the terms

hold for the case of 2 holes. In particular, it is the terms $\mathcal{B}$ and $\mathcal{D}$ that cancel,

$$\mathcal{B} + \mathcal{D} = \quad + \quad = 0, \tag{6.31}$$

and the remaining terms obey,

$$\mathcal{A} + 2\,\mathcal{C} = \quad + 2 \quad = 0. \tag{6.32}$$

This means that the complete VCCD $\mathcal{O}(T^3)$ contribution, $\mathcal{A} + \mathcal{B} + \mathcal{C} + \mathcal{D}$, may again be captured in infinitely many ways, enumerated here by the continuously adjustable parameter μ.

$$\begin{aligned} \mathcal{A} + \mathcal{B} + \mathcal{C} + \mathcal{D} &= \mathcal{A} + \mathcal{C} \\ &= \frac{1}{2}(1-\mu)\mathcal{A} - \mu\,\mathcal{C} \\ &= \frac{1}{2}(1-\mu) \quad - \mu \quad \end{aligned} \tag{6.33}$$

The algebraic expressions for these diagrams in the 2-hole limit are given in Table 6.1. From this, it is quite clear that any method that includes term $\mathcal{B}$ without an appropriately-weighted $\mathcal{D}$ term to cancel it, will be incorrect in this limit. This applies to LPFD(0) and LPFD(-1) in particular. Although AVCCD is correct to $\mathcal{O}(T^3)$, its infinite-order behaviour is based on LPFD(0), and since the cancellations at $\mathcal{O}(T^3)$ are indicative of the cancellations that occur at higher orders, it is similarly incorrect for 2 holes.

Of particular interest, however, is the case $\mu = +1$, which switches off the $\mathcal{A}$ term entirely. In fact, in this limit, $-\mathcal{C}$ correctly accounts for the combined contributions of $\mathcal{A} + \mathcal{B} + \mathcal{C} + \mathcal{D}$, exactly as in the limit of 2 electrons for the case $\lambda = +1$. This is exactly the $\mathcal{O}(T^3)$ contribution generated by LPFD(+1), which implies that, of all the LPFD methods, only LPFD(+1) is correct for 2 holes, at least to $\mathcal{O}(T^3)$.

In fact, a rigorous proof that LPFD(+1) is correct for 2 holes through all orders of the cluster amplitudes, and is therefore exact in this limiting case, can be constructed. The ${}_{\mathcal{C}}\mathbf{U}$ matrix, which is indexed by the unique pairs of electrons, becomes a scalar in the 2-electron limit, since then there is only a

single pair. This element, by construction, takes the value of the CID norm, and, when raised to the appropriate power, introduces division by that norm such that LPFD(+1) is exact in the 2-electron limit. Analogously, it is possible to write down a transformation matrix that is the hole-particle partner of the ${}_{\mathcal{C}}\mathbf{U}$ matrix,

$$ {}_{\mathcal{C}'}\eta^{cd}_{ab} = \frac{1}{2} T^{cd}_{ij} T^{ij}_{ab} \tag{6.34} $$

$$ {}_{\mathcal{C}'}U^{cd}_{ab} = \delta^{cd}_{ab} + \gamma' \, {}_{\mathcal{C}'}\eta^{cd}_{ab}, \tag{6.35} $$

where γ' is taken to be the same as for ${}_{\mathcal{C}}\mathbf{U}$. It would be unwise to use this matrix in any real calculation, however, since the four virtual indices cause it to require extremely large $\mathcal{O}(v^4)$ storage and manipulating it has similarly steep computational complexity. Nevertheless, this matrix, which is indexed by the unique hole pairs, would reduce to a scalar equal to the CID norm in the limit of 2 holes. Thus, using this matrix in place of the ${}_{\mathcal{C}}\mathbf{U}$ matrix in a hypothetical LPFD(+1)-like theory would render the theory exact for 2 holes. Next, consider the binomial series of these matrices when each is raised to an arbitrary rational power, x.

$$ \frac{1}{2} ({}_{\mathcal{C}}\mathbf{U}^x)^{ij}_{kl} \, T^{kl}_{ab} = \frac{1}{2} \sum_{n=0}^{\infty} \binom{x}{n} \gamma'^n \, ({}_{\mathcal{C}}\eta^n)^{ij}_{kl} \, T^{kl}_{ab} \tag{6.36} $$

$$ \frac{1}{2} ({}_{\mathcal{C}'}\mathbf{U}^x)^{cd}_{ab} \, T^{ij}_{cd} = \frac{1}{2} \sum_{n=0}^{\infty} \binom{x}{n} \gamma'^n \, ({}_{\mathcal{C}'}\eta^n)^{cd}_{ab} \, T^{ij}_{cd} \tag{6.37} $$

The following property,

$$ ({}_{\mathcal{C}}\eta^n)^{ij}_{kl} \, T^{kl}_{ab} = ({}_{\mathcal{C}'}\eta^n)^{cd}_{ab} \, T^{ij}_{cd}, \quad n \in \mathbb{N}^0 \tag{6.38} $$

is obviously true for $n = 0$,

$$ \delta^{ij}_{kl} \, T^{kl}_{ab} = \delta^{cd}_{ab} \, T^{ij}_{cd} = T^{ij}_{ab} \tag{6.39} $$

and $n = 1$,

$$ {}_{\mathcal{C}}\eta^{ij}_{kl} \, T^{kl}_{ab} = {}_{\mathcal{C}'}\eta^{cd}_{ab} \, T^{ij}_{cd} = \frac{1}{2} T^{ij}_{cd} T^{cd}_{kl} T^{kl}_{ab} \tag{6.40} $$

and the general result follows by another application of mathematical induction;

assuming the case of $n-1$ to hold, the case of n can be shown to follow.

$$\begin{aligned}
({}_{\mathcal{C}}\eta^{n})^{ij}_{kl}\; T^{kl}_{ab} &= \frac{1}{2}\,{}_{\mathcal{C}}\eta^{ij}_{kl}\left({}_{\mathcal{C}}\eta^{n-1}\right)^{kl}_{mn} T^{mn}_{ab} \\
&= \frac{1}{4}T^{ij}_{cd}\,T^{cd}_{kl}\left({}_{\mathcal{C}}\eta^{n-1}\right)^{kl}_{mn}\; T^{mn}_{ab} \\
&= \frac{1}{4}T^{ij}_{cd}\,T^{cd}_{kl}\left({}_{\mathcal{C}'}\eta^{n-1}\right)^{ef}_{ab}\; T^{kl}_{ef} \\
&= \frac{1}{2}T^{ij}_{cd}\left({}_{\mathcal{C}'}\eta^{n-1}\right)^{ef}_{ab}\; {}_{\mathcal{C}'}\eta^{cd}_{ef} \\
&= ({}_{\mathcal{C}'}\eta^{n})^{cd}_{ab}\; T^{ij}_{cd}
\end{aligned} \tag{6.41}$$

This result proves that the ${}_{\mathcal{C}}\mathbf{U}$ binomial series agrees termwise with the series of the ${}_{\mathcal{C}'}\mathbf{U}$ matrix;

$$\frac{1}{2}({}_{\mathcal{C}}\mathbf{U}^{-\frac{q}{2}})^{ij}_{kl}\,T^{kl}_{ab} = \frac{1}{2}({}_{\mathcal{C}'}\mathbf{U}^{-\frac{q}{2}})^{cd}_{ab}\,T^{ij}_{cd} \tag{6.42}$$

Thus, although the matrices are different in general, they always generate exactly the same VCCD terms through all orders, such that they have exactly the same effect when applied as a transformation matrix to a set of cluster amplitudes. Since the ${}_{\mathcal{C}}\mathbf{U}$ matrix is known to lead to a theory exact for 2 electrons and the ${}_{\mathcal{C}'}\mathbf{U}$ matrix similarly leads to a theory exact for 2 holes, and since both matrices have the same effect when applied to a set of cluster amplitudes, a theory that uses either of these matrices as its principal transformation matrix, such as LPFD(+1), will be exact for *both* 2 electrons and 2 holes.

It is noteworthy, however, that constructing an AVCCD-like method on an LPFD(+1) starting point would yield no better results than the AVCCD theory based on LPFD(0) outlined in the previous chapter. This is because. as has been demonstrated previously, LPFD(+1) performs no better than LPFD(0) in strongly-correlated examples, such as acetylene, and is therefore no better as an underlying theory on which to perform corrections. This was discussed at the end of Chapter 5. The QVCCD ansatz of balancing the contributions to VCCD through all orders of the cluster amplitudes is what is required.

With this result, it is possible to make further progress in identifying constraints among the parameters used to define the QVCCD transformed amplitudes. First, a re-examination of the density matrices for the case of 2 holes, for which $\langle\hat{T}^{\dagger}\hat{T}\rangle = \frac{1}{4}T^{ij}_{ab}T^{ab}_{ij} = \frac{1}{2}T^{ij}_{h\bar{h}}T^{h\bar{h}}_{ij}$, reveals further simplifying relationships to

hold in this limit.

$$\begin{aligned}
{}_{\mathcal{A}}\eta_h^h &= \frac{1}{2}T_{h\bar{h}}^{ij}T_{ij}^{h\bar{h}} = {}_{\mathcal{A}}\eta_{\bar{h}}^{\bar{h}} = \langle\hat{T}^\dagger\hat{T}\rangle, \quad {}_{\mathcal{A}}\eta_{\bar{h}}^{h} = {}_{\mathcal{A}}\eta_h^{\bar{h}} = 0 \\
{}_{\mathcal{B}}\eta_j^i &= T_{h\bar{h}}^{ik}T_{jk}^{h\bar{h}} \\
{}_{\mathcal{C}}\eta_{kl}^{ij} &= T_{h\bar{h}}^{ij}T_{kl}^{h\bar{h}} \quad \text{but} \quad {}_{\mathcal{C}'}\eta_{h\bar{h}}^{h\bar{h}} = \frac{1}{2}T_{ij}^{h\bar{h}}T_{h\bar{h}}^{ij} = \langle\hat{T}^\dagger\hat{T}\rangle \\
{}_{\mathcal{D}}\eta_{hj}^{ih} &= T_{h\bar{h}}^{ik}T_{jk}^{h\bar{h}} = {}_{\mathcal{D}}\eta_{\bar{h}j}^{i\bar{h}} = {}_{\mathcal{B}}\eta_j^i, \quad {}_{\mathcal{D}}\eta_{\bar{h}j}^{ih} = {}_{\mathcal{D}}\eta_{hj}^{i\bar{h}} = 0
\end{aligned} \tag{6.43}$$

By employing the identity above to convert a ${}_{\mathcal{C}}\mathbf{U}$ matrix to a ${}_{\mathcal{C}'}\mathbf{U}$ matrix, it can be shown that both the $\mathcal{A}$ and $\mathcal{C}$ terms can then contribute to the introduction of a CID denominator.

$$\begin{aligned}
{}_q T_{h\bar{h}}^{ij} &\leftarrow \frac{\alpha}{2}(1-\tau_{h\bar{h}})({}_{\mathcal{A}}\mathbf{U}^{{}_{\mathcal{A}}P})_h^c T_{c\bar{h}}^{ij} + \frac{\gamma}{2}({}_{\mathcal{C}}\mathbf{U}^{{}_{\mathcal{C}}P})_{kl}^{ij}T_{h\bar{h}}^{kl} \\
&= \frac{\alpha}{2}(1-\tau_{h\bar{h}})({}_{\mathcal{A}}\mathbf{U}^{{}_{\mathcal{A}}P})_h^h T_{h\bar{h}}^{ij} + \gamma({}_{\mathcal{C}'}\mathbf{U}^{{}_{\mathcal{C}}P})_{h\bar{h}}^{h\bar{h}}T_{h\bar{h}}^{ij} \qquad (6.44) \\
&= \left(\alpha(1+\alpha'\langle\hat{T}^\dagger\hat{T}\rangle)^{{}_{\mathcal{A}}P} + \gamma(1+\gamma'\langle\hat{T}^\dagger\hat{T}\rangle)^{{}_{\mathcal{C}}P}\right)T_{h\bar{h}}^{ij} \qquad (6.45)
\end{aligned}$$

Hence, for the correct introduction of the CID denominator, the following constraints are required,

$${}_{\mathcal{A}}P = {}_{\mathcal{C}}P = -\frac{q}{2} \tag{6.46}$$

$$\alpha' = \gamma' = 1 \tag{6.47}$$

$$\alpha + \gamma = 1, \tag{6.48}$$

and, similarly to the case of 2 electrons, the $\mathcal{B}$ and $\mathcal{D}$ terms can be arranged to cancel in the limit of 2 holes. This follows from the relationship between the ${}_{\mathcal{B}}\eta$ and ${}_{\mathcal{D}}\eta$ matrices that again can be shown to generalize to integer powers of these matrices by a final application of mathematical induction; the following property,

$$({}_{\mathcal{D}}\eta^n)_{hj}^{ih} = ({}_{\mathcal{D}}\eta^n)_{\bar{h}j}^{i\bar{h}} = ({}_{\mathcal{B}}\eta^n)_j^i\,, \quad n \in \mathbb{N}^0 \tag{6.49}$$

has already been shown to hold for $n = 1$, and trivially holds for $n = 0$,

$$\left({}_{\mathcal{D}}\eta^0\right)_{hj}^{ih} = \delta_{hj}^{ih} = \delta_j^i = \left({}_{\mathcal{B}}\eta^0\right)_j^i\,. \tag{6.50}$$

Then, assuming it to hold for all integers up to n allows the case of $n+1$ to be

proved, completing the inductive step.

$$\begin{aligned}
\left({}_{\mathcal{D}}\eta^{n+1}\right)^{ih}_{hj} &= \left({}_{\mathcal{D}}\eta\right)^{ic}_{hk}\left({}_{\mathcal{D}}\eta^{n}\right)^{kh}_{cj} \\
&= \left({}_{\mathcal{D}}\eta\right)^{ih}_{hk}\left({}_{\mathcal{D}}\eta^{n}\right)^{kh}_{hj} \\
&= \left({}_{\mathcal{B}}\eta\right)^{i}_{k}\left({}_{\mathcal{B}}\eta^{n}\right)^{k}_{j} \\
&= \left({}_{\mathcal{B}}\eta^{n+1}\right)^{i}_{j} \quad n \in \mathbb{N}^0
\end{aligned} \tag{6.51}$$

Using this result, the conditions for the appropriate cancellation of the $\mathcal{B}$ and $\mathcal{D}$ terms can be found.

$$\begin{aligned}
{}_{q}T^{ij}_{h\bar{h}} \leftarrow{}& \frac{1}{2}\beta(1-\tau_{ij})\left({}_{\mathcal{B}}\mathbf{U}^{\,{}_{\mathcal{B}}P}\right)^{i}_{k}T^{kj}_{h\bar{h}} + \frac{1}{4}\delta(1-\tau_{ij})(1-\tau_{h\bar{h}})\left({}_{\mathcal{D}}\mathbf{U}^{\,{}_{\mathcal{D}}P}\right)^{ih}_{hk}T^{kj}_{h\bar{h}} \\
={}& \sum_{n=0}^{\infty}\binom{{}_{\mathcal{B}}P}{n}\left[\frac{1}{2}\beta(1-\tau_{ij})\left(\beta'^{n}\,{}_{\mathcal{B}}\eta^{n}\right)^{i}_{k}T^{kj}_{h\bar{h}}\right] \\
&+ \sum_{n=0}^{\infty}\binom{{}_{\mathcal{D}}P}{n}\left[\frac{1}{4}\delta(1-\tau_{ij})(1-\tau_{h\bar{h}})\left(\delta'^{n}\,{}_{\mathcal{D}}\eta^{n}\right)^{ih}_{hk}T^{kj}_{h\bar{h}}\right] \\
={}& \sum_{n=0}^{\infty}\binom{{}_{\mathcal{B}}P}{n}\left[\frac{1}{2}\beta(1-\tau_{ij})\left(\beta'^{n}\,{}_{\mathcal{B}}\eta^{n}\right)^{i}_{k}T^{kj}_{h\bar{h}}\right] \\
&+ \sum_{n=0}^{\infty}\binom{{}_{\mathcal{D}}P}{n}\left[\frac{1}{4}\delta(1-\tau_{ij})(1-\tau_{h\bar{h}})\left(\delta'^{n}\,{}_{\mathcal{B}}\eta^{n}\right)^{i}_{k}T^{kj}_{h\bar{h}}\right] \\
={}& \frac{1}{2}\sum_{n=0}^{\infty}\binom{{}_{\mathcal{B}}P}{n}\left[\beta(1-\tau_{ij})\left(\beta'^{n}\,{}_{\mathcal{B}}\eta^{n}\right)^{i}_{k}T^{kj}_{h\bar{h}}\right] \\
&+ \frac{1}{2}\sum_{n=0}^{\infty}\binom{{}_{\mathcal{D}}P}{n}\left[\delta(1-\tau_{ij})\left(\delta'^{n}\,{}_{\mathcal{B}}\eta^{n}\right)^{i}_{k}T^{kj}_{h\bar{h}}\right] \\
={}& \frac{1}{2}\beta\sum_{n=0}^{\infty}\binom{{}_{\mathcal{B}}P}{n}\beta'^{n}(1-\tau_{ij})\left({}_{\mathcal{B}}\eta^{n}\right)^{i}_{k}T^{kj}_{h\bar{h}} \\
&+ \frac{1}{2}\delta\sum_{n=0}^{\infty}\binom{{}_{\mathcal{D}}P}{n}\delta'^{n}(1-\tau_{ij})\left({}_{\mathcal{B}}\eta^{n}\right)^{i}_{k}T^{kj}_{h\bar{h}} \\
={}& \frac{1}{2}\sum_{n=0}^{\infty}\left[\beta\,\beta'^{n}\binom{{}_{\mathcal{B}}P}{n} + \delta\,\delta'^{n}\binom{{}_{\mathcal{D}}P}{n}\right](1-\tau_{ij})\left({}_{\mathcal{B}}\eta^{n}\right)^{i}_{k}T^{kj}_{h\bar{h}}
\end{aligned} \tag{6.52}$$

Again discarding trivial cases such as $\beta = \delta = 0$, which would permanently switch off the corresponding transformations, this clearly imposes the following constraints for the cancellation to occur as required,

$${}_{\mathcal{B}}P - {}_{\mathcal{D}}P = 0 \tag{6.53}$$

$$\beta' - \delta' = 0 \tag{6.54}$$

$$\beta + \delta = 0. \tag{6.55}$$

The cancellations that have been shown to be necessary so far in this chapter mimic those that were noted earlier to occur between the $\mathcal{O}(T^3)$ VCCD terms in the limiting cases of either 2 electrons or 2 holes; in the case of 2 electrons, terms $\mathcal{A}$ and $\mathcal{D}$ cancel exactly, whereas for 2 holes, terms $\mathcal{B}$ and $\mathcal{D}$ cancel exactly. The algebraic expressions to which these terms correspond are given in Table 6.1.

6.4 Derivation III - Constraints Arising from the Third-Order VCCD Terms

At this point it is useful to collate all of the constraints derived so far. First, the following relationships have been derived for the powers,

$$_{\mathcal{B}}P = {}_{\mathcal{C}}P = -\frac{q}{2} \tag{6.56}$$

$$_{\mathcal{A}}P - {}_{\mathcal{D}}P = 0 \tag{6.57}$$

$$_{\mathcal{A}}P = {}_{\mathcal{C}}P = -\frac{q}{2} \tag{6.58}$$

$$_{\mathcal{B}}P - {}_{\mathcal{D}}P = 0. \tag{6.59}$$

Thus, the values of these parameters are already fully determined.

$$_{\mathcal{A}}P = {}_{\mathcal{B}}P = {}_{\mathcal{C}}P = {}_{\mathcal{D}}P = -\frac{q}{2} \tag{6.60}$$

Similarly, the following constraints have been derived for the coefficients of the density matrices,

$$\beta' = \gamma' = 1 \tag{6.61}$$

$$\alpha' - \delta' = 0 \tag{6.62}$$

$$\alpha' = \gamma' = 1 \tag{6.63}$$

$$\beta' - \delta' = 0, \tag{6.64}$$

and so these parameters are also already fully determined to be,

$$\alpha' = \beta' = \gamma' = \delta' = 1. \tag{6.65}$$

Therefore, eight of the twelve parameters have been fixed from considerations of the limiting cases of 2 electrons and 2 holes, and all that remains are the

coefficients of the transformations, for which the following constraints have been derived,

$$\alpha + \beta + \gamma + \delta = 1 \tag{6.66}$$
$$\beta + \gamma = 1 \tag{6.67}$$
$$\alpha + \delta = 0 \tag{6.68}$$
$$\alpha + \gamma = 1 \tag{6.69}$$
$$\beta + \delta = 0. \tag{6.70}$$

Note that second and third constraints can be added together to yield the first constraint, and the same can be done with the fourth and fifth. Thus, there are only three unique constraints currently known,

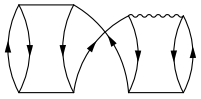
$$\alpha + \beta + \gamma + \delta = 1 \tag{6.71}$$
$$\beta + \gamma = 1 \tag{6.72}$$
$$\alpha + \gamma = 1, \tag{6.73}$$

which is insufficient to uniquely determine the values of these four parameters, and further information must be sought. This information can be found quite readily from an examination of the $\mathcal{O}(T^3)$ VCCD terms generated by each of the transformations.

A detailed breakdown of the Coupled Cluster diagrams and corresponding terms that contribute to $\frac{1}{2!}\langle \hat{T}^\dagger \hat{H} \hat{T}^2 \rangle_L$ is given in Table 6.1, in which the values of the terms for the limit of 2 holes is also given. With this information at hand, it is possible to readily determine the weightings of each of the $\mathcal{O}(T^3)$ terms relative to how they appear in VCCD. These relative weightings should, of course, all be made equal to 1 if QVCCD is to match VCCD exactly to $\mathcal{O}(T^3)$. For example,

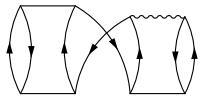
the $\mathcal{A}$ term is generated with a factor of $+\alpha/2$,

$$\begin{aligned}
\frac{1}{4}\langle ij||ab\rangle\, {}_2T_{ij}^{ab} &\leftarrow \frac{1}{4}\langle ij||ab\rangle.\alpha\left[\frac{1}{2}(1-\tau_{ab})\left({}_{\mathcal{A}}\mathbf{U}^{-1}\right)_c^a T_{ij}^{cb}\right] \\
&= \frac{1}{8}\alpha\langle ij||ab\rangle(1-\tau_{ab})\sum_{n=0}^{\infty}\binom{-1}{n}\left({}_{\mathcal{A}}\eta^n\right)_c^a T_{ij}^{cb} \\
&\leftarrow -\frac{1}{8}\alpha\langle ij||ab\rangle(1-\tau_{ab})\,{}_{\mathcal{A}}\eta_c^a T_{ij}^{cb} \\
&= -\frac{1}{4}\alpha\langle ij||ab\rangle\,{}_{\mathcal{A}}\eta_c^a T_{ij}^{cb} \\
&= -\frac{1}{4}\alpha\langle ij||ab\rangle\,{}_{\mathcal{A}}\eta_c^b T_{ij}^{ac} \\
&= +\frac{\alpha}{2}
\end{aligned} \qquad (6.74)$$



Given that the $\mathcal{B}$ term is the hole-particle opposite of the $\mathcal{A}$ term, it makes intuitive sense that it is generated with an analogous factor of $+\beta/2$. In fact, this agrees with the terms generated by LPFD(0), for which $\alpha = \gamma = \delta = 0$ and $\beta = +1$, and which generates $\mathcal{B}$ with a factor of $1/2$.

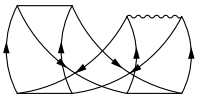
$$\begin{aligned}
\frac{1}{4}\langle ij||ab\rangle\, {}_2T_{ij}^{ab} &\leftarrow \frac{1}{4}\langle ij||ab\rangle.\beta\left[\frac{1}{2}(1-\tau_{ij})\left({}_{\mathcal{B}}\mathbf{U}^{-1}\right)_i^k T_{kj}^{ab}\right] \\
&= \frac{1}{8}\beta\langle ij||ab\rangle(1-\tau_{ij})\sum_{n=0}^{\infty}\binom{-1}{n}\left({}_{\mathcal{B}}\eta^n\right)_i^k T_{kj}^{ab} \\
&\leftarrow -\frac{1}{8}\beta\langle ij||ab\rangle(1-\tau_{ij})\,{}_{\mathcal{B}}\eta_i^k T_{kj}^{ab} \\
&= -\frac{1}{4}\beta\langle ij||ab\rangle\,{}_{\mathcal{B}}\eta_i^k T_{kj}^{ab} \\
&= -\frac{1}{4}\beta\langle ij||ab\rangle\,{}_{\mathcal{B}}\eta_j^k T_{ik}^{ab} \\
&= +\frac{\beta}{2}
\end{aligned} \qquad (6.75)$$



The $\mathcal{C}$ term is generated with a factor of $-\gamma$, which agrees with the terms present in LPFD(+1), for which $\alpha = \beta = \delta = 0$ and $\gamma = +1$, and for which the $\mathcal{C}$ term is

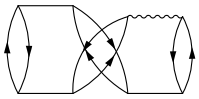
generated with a factor of -1.

$$\begin{aligned}
\frac{1}{4}\langle ij||ab\rangle\, {}_2T_{ij}^{ab} &\leftarrow \frac{1}{4}\langle ij||ab\rangle.\gamma\left[\frac{1}{2}\left({}_{\mathcal{C}}\mathbf{U}^{-1}\right)_{ij}^{kl} T_{kl}^{ab}\right] \\
&= \frac{1}{8}\gamma\langle ij||ab\rangle \sum_{n=0}^{\infty}\binom{-1}{n}({}_{\mathcal{C}}\eta^{n})_{ij}^{kl} T_{kl}^{ab} \\
&\leftarrow -\frac{1}{8}\gamma\langle ij||ab\rangle\, {}_{\mathcal{C}}\eta_{ij}^{kl} T_{kl}^{ab} \\
&= -\gamma \quad \text{(6.76)}
\end{aligned}$$



The final linked $\mathcal{O}(T^3)$ VCCD, term $\mathcal{D}$, can be shown to be generated with a factor of $-\delta/2$,

$$\begin{aligned}
\frac{1}{4}\langle ij||ab\rangle\, {}_2T_{ij}^{ab} &\leftarrow \frac{1}{4}\langle ij||ab\rangle.\delta\left[\frac{1}{4}(1-\tau_{ij})(1-\tau_{ab})\left({}_{\mathcal{D}}\mathbf{U}^{-1}\right)_{ic}^{ak} T_{kj}^{cb}\right] \\
&= \frac{1}{16}\delta\langle ij||ab\rangle(1-\tau_{ij})(1-\tau_{ab})\sum_{n=0}^{\infty}\binom{-1}{n}({}_{\mathcal{D}}\eta^{n})_{ic}^{ak} T_{kj}^{cb} \\
&\leftarrow -\frac{1}{16}\delta\langle ij||ab\rangle(1-\tau_{ij})(1-\tau_{ab})\, {}_{\mathcal{D}}\eta_{ic}^{ak} T_{kj}^{cb} \\
&= -\frac{1}{4}\delta\langle ij||ab\rangle\, {}_{\mathcal{D}}\eta_{ic}^{ak} T_{kj}^{cb} \\
&= -\frac{1}{4}\delta\langle ij||ab\rangle T_{il}^{ad} T_{dc}^{lk} T_{kj}^{cb} \\
&= -\frac{1}{4}\delta\langle ij||ab\rangle T_{ik}^{ac} T_{cd}^{kl} T_{lj}^{db} \\
&= -\frac{\delta}{2} \quad \text{(6.77)}
\end{aligned}$$

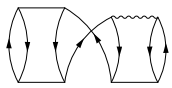
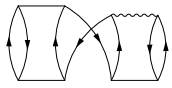
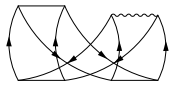
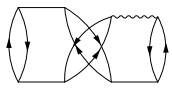


Thus, with all the weightings of the $\mathcal{O}(T^3)$ terms determined as functions of the remaining unknown parameters, the values of these parameters can be determined through the requirement that each weighting be made equal to $+1$.

$$\begin{aligned}
+\frac{\alpha}{2} = +1 &\implies \alpha = +2 \\
+\frac{\beta}{2} = +1 &\implies \beta = +2 \\
-\gamma = +1 &\implies \gamma = -1 \\
-\frac{\delta}{2} = +1 &\implies \delta = -2 \quad \text{(6.78)}
\end{aligned}$$

Finally, it must be checked that these values are consistent with the constraints given in Equations 6.71-6.73 that were derived previously, such that, with these

Table 6.1: Linked $\mathcal{O}(T^3)$ contributions to VCCD, where $\langle \hat{T}^\dagger \hat{T} \rangle = \frac{1}{4} T_{ab}^{ij} T_{ij}^{ab}$, and where e and $\bar{e}$ and h and $\bar{h}$ label the two spinorbitals occupied and unoccupied in the reference wavefunction respectively.

	$\frac{1}{2!}\langle \hat{T}_2^\dagger \hat{H} \hat{T}_2^2 \rangle_L$ Diagram	$\frac{1}{2!}\langle \hat{T}_2^\dagger \hat{H} \hat{T}_2^2 \rangle_L$ Term	2 Electrons	2 Holes
$\mathcal{A}$		$\frac{1}{4} T_{ij}^{ac} T_{cd}^{kl} T_{kl}^{db} \langle ij \vert\vert ab \rangle$ $= -\frac{1}{2} T_{ij}^{ac} \, {}_{\mathcal{A}}\eta_c^b \langle ij \vert\vert ab \rangle$	$-T_{e\bar{e}}^{ac} \, {}_{\mathcal{A}}\eta_c^b \langle e\bar{e} \vert\vert ab \rangle$	$-T_{ij}^{h\bar{h}} \langle \hat{T}^\dagger \hat{T} \rangle \langle ij \vert\vert h\bar{h} \rangle$
$\mathcal{B}$		$\frac{1}{4} T_{ik}^{ab} T_{cd}^{kl} T_{lj}^{cd} \langle ij \vert\vert ab \rangle$ $= -\frac{1}{2} T_{ik}^{ab} \, {}_{\mathcal{B}}\eta_j^k \langle ij \vert\vert ab \rangle$	$-T_{e\bar{e}}^{ab} \langle \hat{T}^\dagger \hat{T} \rangle \langle e\bar{e} \vert\vert ab \rangle$	$-T_{ik}^{h\bar{h}} \, {}_{\mathcal{B}}\eta_j^k \langle ij \vert\vert h\bar{h} \rangle$
$\mathcal{C}$		$\frac{1}{16} T_{kl}^{ab} T_{cd}^{kl} T_{ij}^{cd} \langle ij \vert\vert ab \rangle$ $= \frac{1}{8} T_{kl}^{ab} \, {}_{\mathcal{C}}\eta_{ij}^{kl} \langle ij \vert\vert ab \rangle$ $= \frac{1}{8} T_{ij}^{cd} \, {}_{\mathcal{C}}\eta_{cd}^{ab} \langle ij \vert\vert ab \rangle$	$\frac{1}{2} T_{e\bar{e}}^{ab} \langle \hat{T}^\dagger \hat{T} \rangle \langle e\bar{e} \vert\vert ab \rangle$	$\frac{1}{2} T_{ij}^{h\bar{h}} \langle \hat{T}^\dagger \hat{T} \rangle \langle ij \vert\vert h\bar{h} \rangle$
$\mathcal{D}$		$\frac{1}{2} T_{ik}^{ac} T_{cd}^{kl} T_{lj}^{db} \langle ij \vert\vert ab \rangle$ $= \frac{1}{2} T_{ik}^{ac} \, {}_{\mathcal{D}}\eta_{cj}^{kb} \langle ij \vert\vert ab \rangle$	$T_{e\bar{e}}^{ac} \, {}_{\mathcal{A}}\eta_c^b \langle e\bar{e} \vert\vert ab \rangle$	$T_{ik}^{h\bar{h}} \, {}_{\mathcal{B}}\eta_j^k \langle ij \vert\vert h\bar{h} \rangle$

values, the QVCCD method treats the limiting cases of 2 electrons and 2 holes exactly and agrees with VCCD to $\mathcal{O}(T^3)$. Fortunately, these values pass this test.

$$\alpha + \beta + \gamma + \delta = 2 + 2 - 1 - 2 = 1 \quad \checkmark$$
$$\beta + \gamma = 2 - 1 = 1 \quad \checkmark$$
$$\alpha + \gamma = 2 - 1 = 1 \quad \checkmark$$

6.5 The Fully-Determined Quasi-Variational Coupled Cluster Method

In the last three sections, all twelve of the parameters defining the QVCCD method have been determined uniquely from the requirements that the functional correspond to CID for limiting systems consisting of 2 electrons or 2 holes and from the correct reproduction of the $\mathcal{O}(T^3)$ VCCD terms for general systems. In contrast to both LPFD, which contained the arbitrary parameter λ, and AVCCD, which additionally contained arbitrary coefficients and powers used in the corrective transformations, QVCCD is a truly *ab initio* theory. This new quantum-chemical method will be the primary focus of this thesis from this point on. Since it has such central importance, and since its derivation has been dis-

cussed in such detail, a summary of the workings and properties of this theory is given below.

The QVCCD method is defined through the following functional,

$$E_{\mathrm{QVCCD}} = \langle \hat{H} \rangle + 2\langle \hat{H}\, {}_2\hat{T} \rangle + \langle {}_1\hat{T}^{\dagger}(\hat{H} - \langle \hat{H} \rangle)\, {}_1\hat{T} \rangle, \tag{6.79}$$

and the ground-state QVCCD energy corresponds to the minimum of this functional with respect to the set of doubles-only cluster amplitudes, $\{T_{ab}^{ij}\}$. The transformed cluster amplitudes take the form of a linear combination of several different LPF-style amplitude transformations, which seeks to achieve a balance of several types of VCCD terms through all orders of the cluster amplitudes.

$$\begin{aligned}
{}_q\hat{T}|\Phi_0\rangle &= \frac{1}{4}\, {}_qT_{ab}^{ij}|\Phi_{ij}^{ab}\rangle \\
{}_{\mathcal{A}}U_b^a &= \delta_b^a + \frac{1}{2}T_{bc}^{ij}\, T_{ij}^{ac} \\
{}_{\mathcal{B}}U_j^i &= \delta_j^i + \frac{1}{2}T_{ab}^{ik}\, T_{jk}^{ab} \\
{}_{\mathcal{C}}U_{kl}^{ij} &= \delta_{kl}^{ij} + \frac{1}{2}T_{ab}^{ij}\, T_{kl}^{ab} \\
{}_{\mathcal{D}}U_{aj}^{ib} &= \delta_{aj}^{ib} + T_{ac}^{ik}\, T_{jk}^{bc}
\end{aligned}
\qquad
\begin{aligned}
{}_qT_{ab}^{ij} = &+ 2\left[\frac{1}{2}(1-\tau_{ab})\left({}_{\mathcal{A}}\mathbf{U}^{-\frac{q}{2}}\right)_a^c T_{cb}^{ij}\right] \\
&+ 2\left[\frac{1}{2}(1-\tau_{ij})\left({}_{\mathcal{B}}\mathbf{U}^{-\frac{q}{2}}\right)_k^i T_{ab}^{kj}\right] \\
&- 1\left[\frac{1}{2}\left({}_{\mathcal{C}}\mathbf{U}^{-\frac{q}{2}}\right)_{kl}^{ij} T_{ab}^{kl}\right] \\
&- 2\left[\frac{1}{4}(1-\tau_{ij})(1-\tau_{ab})\left({}_{\mathcal{D}}\mathbf{U}^{-\frac{q}{2}}\right)_{ak}^{ic} T_{cb}^{kj}\right]
\end{aligned}
\tag{6.80}$$

When this functional is also optimized with respect to the orbitals[100] (OQVCCD), or constrained by a Brueckner condition[96–98] (BQVCCD), in order to account for the effects of single excitations, the ansatz possesses the following extremely theoretically attractive set of methodological properties.

- OQVCCD is a unique theory, containing no arbitrary parameters.
- The ground-state energy is calculated by variational minimization of a functional. The error in a calculated energy is thus second-order in any remaining errors in the cluster amplitude parameters.
- The functional contains fully linked terms only, and is therefore rigorously extensive.
- It is exactly equivalent to FCI for 2 electrons or 2 holes, systems for which the cluster operator $\hat{T} = \hat{T}_1 + \hat{T}_2$ is complete.

- The energy is a scalar that is invariant to rotations in the underlying orbital spaces $\{\psi_i\}$ and $\{\psi_a\}$.
- The energy is not a true upper bound on the exact ground-state Schrödinger energy eigenvalue, but variational minimization of the functional gives a theory that satisfies the generalized Hellmann-Feynman theorem[67].
- The doubles-only theory is equivalent to VCCD to $\mathcal{O}(T^3)$, and, as a consequence, the one-electron $\mathcal{O}(T^4)$ terms are also constructed correctly.
- Furthermore, by construction, balanced VCCD contributions are generated through all orders of the cluster amplitudes, such that the functional potentially inherits approximate upper bound character from its parent VCCD method.
- It is correct to third-order in Møller-Plesset perturbation theory and omits only the terms containing triple excitations from fourth-order, the same as CCSD.
- The limiting computational complexity of the method is $\mathcal{O}(o^2v^4)$ with $\mathcal{O}(o^2v^2)$ storage, the same as CCSD.

The remarks made about the one-electron $\mathcal{O}(T^4)$ terms and the relationship of the method to Møller-Plesset perturbation theory will be discussed in the following chapter, but a brief discussion of the computational complexity is in order here. The overall computational cost of performing an OQVCCD calculation is very slightly higher than performing an equivalent CCSD calculation. This is partially because the use of optimization of the orbitals to treat the single excitations requires an integral transformation for each iteration of the program. This involves up to three virtual orbitals, at $\mathcal{O}(o^3v^3)$ cost. For technical reasons, the integral transformation associated with the Brueckner condition is computationally simpler, requiring only $\mathcal{O}(o^4v^2)$ time, but still introduces additional steps in comparison to the CCSD method. Furthermore, the minimization of the QVCCD functional requires the solution of the eigenproblems for each of the four matrices, ${}_{\mathcal{A}}\mathbf{U}$, ${}_{\mathcal{B}}\mathbf{U}$, ${}_{\mathcal{C}}\mathbf{U}$ and ${}_{\mathcal{D}}\mathbf{U}$, for reasons outlined in Section 4.1, on each iteration of the program. Since the solution of the eigenproblem for an $N \times N$ matrix

requires $\mathcal{O}(N^3)$ time, the complexity of each of these eigenproblems is $\mathcal{O}(v^3)$, $\mathcal{O}(o^3)$, $\mathcal{O}(o^6)$ and $\mathcal{O}(o^3v^3)$ respectively. Of these, it is only the $\mathcal{O}(o^3v^3)$ steps involved for ${}_{\mathcal{D}}\mathbf{U}$ that are concerning, and future work should address whether a finite truncation of the binomial series of ${}_{\mathcal{D}}\mathbf{U}^{-\frac{q}{2}}$ is possible, either in principle or in practice. However, the manipulation of the four-external integrals in either CCSD or OQVCCD requires $\mathcal{O}(o^2v^4)$ time, and this is thus the limiting step in both calculations. Therefore, it should be true that if it is computationally practical to perform a CCSD calculation on a particular system, it should also be possible to perform an equivalent OQVCCD calculation.

6.6 Preliminary Results

In this, the final section of this chapter, some preliminary results are given and discussed in order to establish the adequacy of OQVCCD theory as an approximation to VCCSD. As was noted in the previous chapter, the earlier OAVCCD method predicts potential energy curves qualitatively more like VCCSD than CCSD for the three systems tested; BeO, H_4 and C_2H_2. These systems are sufficiently complicated that the CCSD method struggles to perform adequately, diverging from the VCCSD method to yield results qualitatively different from FCI and sometimes physically incorrect altogether. Despite this, OAVCCD is not itself particularly quantitatively accurate. These systems must be tested again with the OQVCCD method in order to determine if it yields improved results, and whether the improved approximation of VCCSD allows the OQVCCD method to inherit sufficient upper bound character to treat difficult problems such as breaking the triple bond in acetylene.

The new plot for BeO, showing the additional potential energy curve predicted by OQVCCD, is shown in Fig. 6.3. As was previously noted, the OAVCCD method is in excellent agreement with VCCSD in the interval 1.5-2.5Å, and the OQVCCD method performs similarly well. This is in contrast to CCSD, which experiences a crossing with the FCI curve at approximately 2.2Å, and BCCD which predicts an unphysical maximum. However, the OAVCCD method does not describe the dissociation quantitatively accurately, possessing a large error

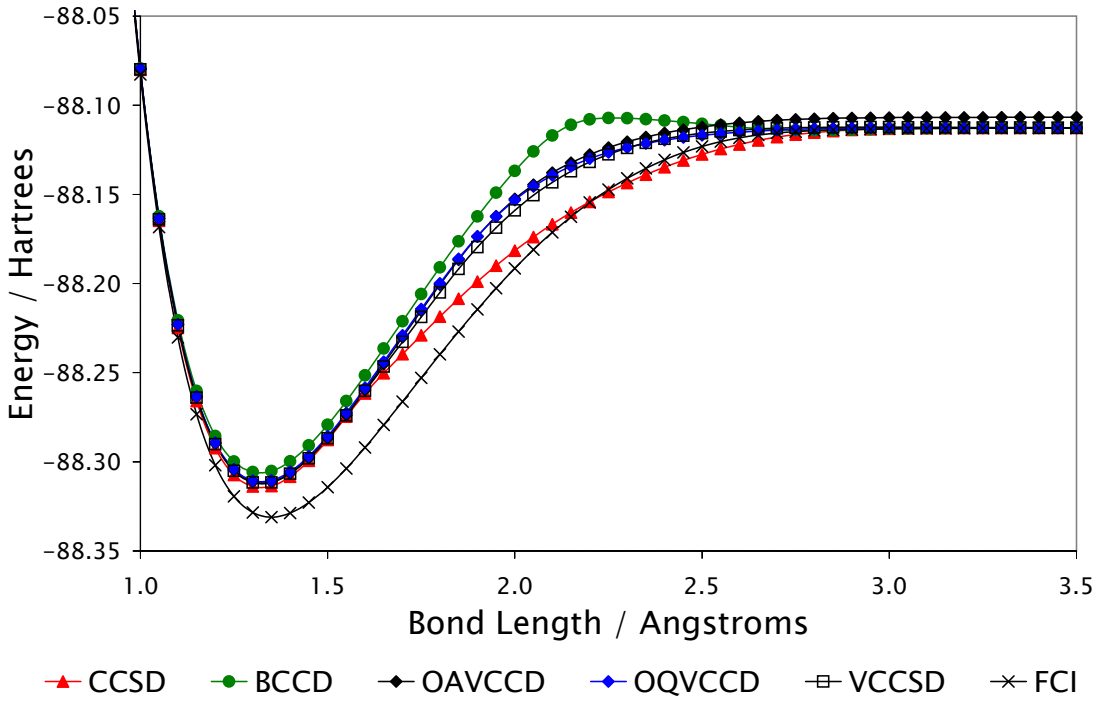


Figure 6.3: Calculated potential energy curves for BeO with the STO-3G basis set.

relative to both VCCSD and FCI from approximately 2.5Å. Even the TCC methods converge to the correct limit at long bond lengths in this system, making this behaviour problematic for OAVCCD. In contrast, not only does OQVCCD describe the intermediate interval with an accuracy that matches or exceeds that of OAVCCD, the predicted potential energy curve also levels off much more quickly than OAVCCD, and approaches the correct dissociation limit.

Similarly, the OQVCCD method predicts a significantly improved cusp for the H_4 model system. This is illustrated in Fig. 6.4. Previously, it was noted that both CCSD and BCCD erroneously predict an inverted cusp that is at odds with FCI, for which a smooth maximum is present at $\theta = 90°$, but, while VCCSD still predicts a cusp, it is a much better approximation to the FCI curve. The OAVCCD method predicts a cusp that is qualitatively like that of VCCSD, but is far too high in energy, and is thus itself a poor model of the FCI behaviour. The OQVCCD curve however, is almost coincident with the VCCSD curve throughout. Since dynamic correlation is weak in this system, but static correlation is strong due to the swapping of the optimum references around $\theta = 90°$, this provides compelling evidence that OQVCCD is capable of treating static correlation

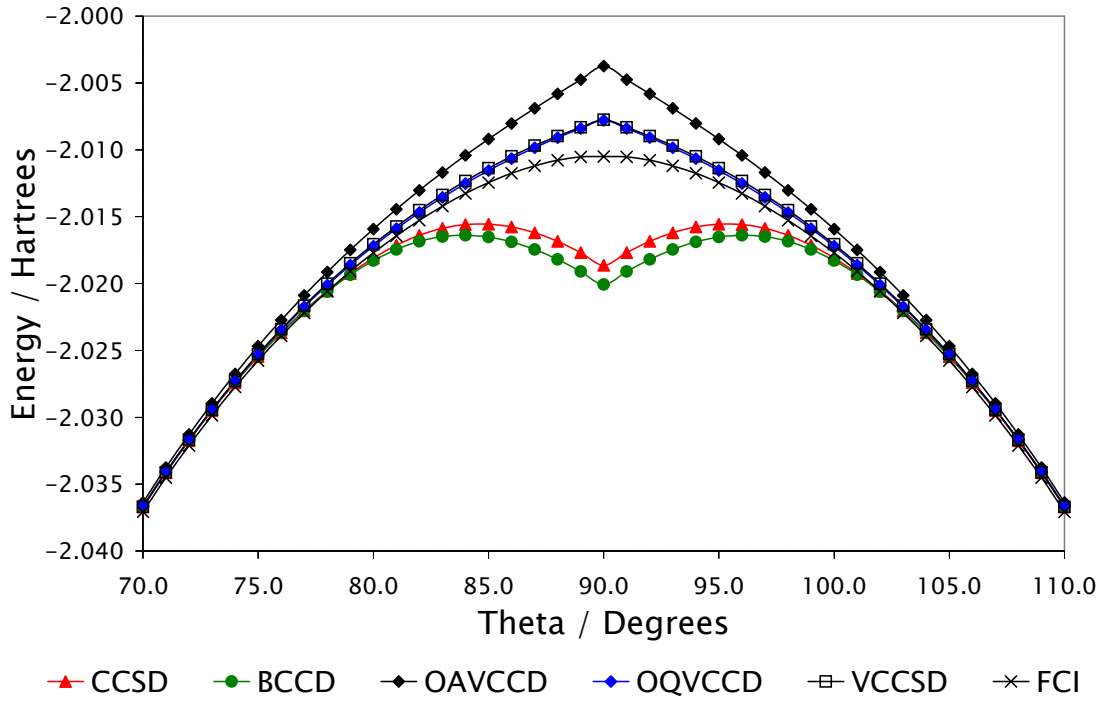


Figure 6.4: Calculated potential energy curves for H_4 with $R = 1.75$Å, and with the aug-cc-pVDZ basis set.

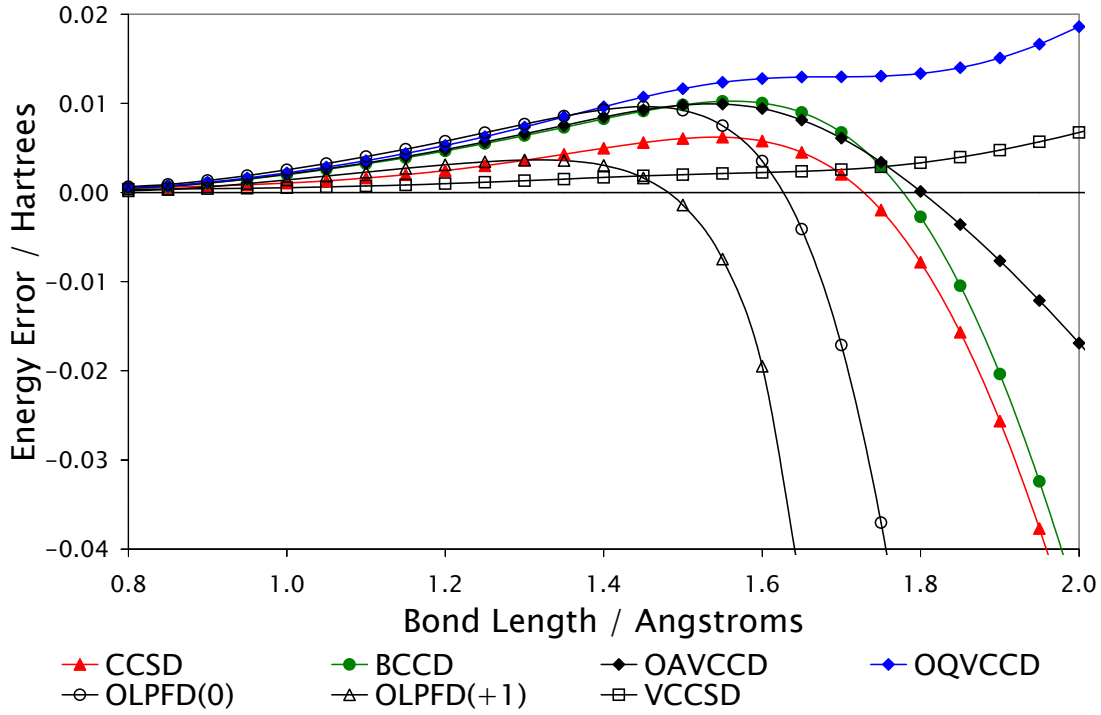


Figure 6.5: Errors relative to FCI in calculated energies for the stretching of the carbon-carbon triple bond in acetylene, C_2H_2, with the C-H bond length fixed at 1.06Å, and with the STO-3G basis set.

with a quality similar to VCCSD.

Finally, potential energy curves for breaking the triple bond in acetylene are re-examined. For clarity, the plot in Fig. 6.5 is again presented as errors relative to FCI. As is obvious from the figure, CCSD, BCCD, OLPFD(0), OLPFD(+1) and OAVCCD all perform non-variationally, crossing the FCI potential energy curve as the triple bond is stretched. VCCSD does not experience a crossing due to its upper bound property, and it is extremely promising that OQVCCD also remains above the FCI energy; despite a significantly larger error than VCCSD at long bond lengths, OQVCCD remains physically correct throughout. Each of the impressive results obtained for these three systems prompt the benchmarking of the OQVCCD method on further systems and with more representative basis sets. This will be performed in the next chapter, once a perturbative correction for the effects of triples has been obtained for the OQVCCD method.

Chapter 7

Perturbative Corrections for Triples and Benchmark Results[†]

The previous chapter detailed the construction of the OQVCCD approximation, and presented preliminary results of calculations that indicate it to inherit significant upper bound character from its parent theory, VCCSD, such that it is more robust to the breakdown of the Hartree-Fock approximation than methods based on Traditional Coupled Cluster theory. In particular, the OQVCCD method appears to possess enough upper bound character to treat difficult problems such as the breaking of multiple bonds, for example in acetylene, sufficiently well that predicted potential energy curves remain physically correct throughout. It is the purpose of this chapter to more thoroughly benchmark the OQVCCD method against TCC-based methods, especially on multiply-bonded systems, for which it was designed.

However, it is well known that a method restricted only to single and double excitations of a single-determinantal reference wavefunction captures insufficient dynamic electron correlation to achieve chemically accurate results, and that the effect of at least triple excitations must be additionally included. Unfortunately,

[†]Relevant publications:

[104] J. B. Robinson and P. J. Knowles, Phys. Chem. Chem. Phys. **14**, 6729 (2012).

[105] J. B. Robinson and P. J. Knowles, J. Chem. Theory Comput. (2012), *"Benchmark Quasi-Variational Coupled Cluster Calculations of Multiple Bond Breaking"*, in press.

the explicit inclusion of connected triples is prohibitively expensive. For example, while the CCSD method scales in computational complexity as $\mathcal{O}(N^6)$, where N is some measure of the size of the system, such as the number of electrons, the CCSDT method scales as $\mathcal{O}(N^8)$. The CCSDTQ method is even more computationally expensive, possessing $\mathcal{O}(N^{10})$ complexity. For this reason, the effects of connected triple excitations are commonly included approximately (through perturbation theory) in a single non-iterative $\mathcal{O}(N^7)$ step.

In the case of CCSD, numerous such perturbative corrections have been proposed. Analysing the contributions to CCSD through orders of Møller-Plesset perturbation theory, CCSD first diverges from FCC at fourth-order, since it omits those terms that contain triple excitations. The [T] correction[106] is the minimal non-iterative correction for these missing terms, such that CCSD[T] is correct to fourth-order. The CCSD(T)[107] method also includes some additional terms at fifth-order and higher that have been justified in different ways[70, 107, 108], and further corrective terms have also been proposed[109].

However, it is also well-known[110] that CCSD(T) fails to accurately describe molecular dissociation due to the breakdown of the (T) perturbative correction, which can become singular if the Highest Occupied Molecular Orbital (HOMO) and Lowest Unoccupied Molecular Orbital (LUMO) become degenerate, common when bonds are stretched. To remedy this failure of CCSD(T) for bond-breaking situations, several authors have proposed more advanced alternatives, such as the Λ methods of Bartlett et al.[83], the CCSD(2) method[111] and the Completely Renormalized CC (CR-CC(2,3)) method of Piecuch et al.[112].

In this chapter, the perturbative decompositions of TCC and VCC are discussed. For simplicity, single excitations are taken to vanish, as in the case of using Brueckner or variationally optimal orbitals. These perturbative decompositions allow comparisons between the corrections necessary for TCC and for VCC to be drawn, allowing the minimalistic [T] correction for the omission of connected triple excitations by VCCD, and subsequently by QVCCD, to be determined. This is then generalized to (T). The resulting OQVCCD(T) method is then benchmarked for both single-bond dissociations, and also multiple-bond dissociations for which non-dynamic correlation is especially strong, and for which

CCSD(T) is known to fail catastrophically, with some surprising results.

7.1 The Perturbative Decomposition of TCC and VCC

As is common in Møller-Plesset analyses, the Hamiltonian is partitioned into the one-electron Fock operator, $\hat{H}^{(0)} = \hat{F}$, and the 2-electron fluctuation potential, $\hat{H}^{(1)} = \hat{V}$, such that $\hat{H} = \hat{F} + \hat{V}$. With the Hartree-Fock reference wavefunction taken to be the zeroth-order wavefunction (that is Hartree-Fock is assumed to be the problem that is solved exactly), the energy may then be decomposed in orders of perturbation theory,

$$E = E^{(0)} + E^{(1)} + \ldots, \tag{7.1}$$

along with the cluster operators,

$$\hat{T} = \hat{T}^{(1)} + \hat{T}^{(2)} + \ldots . \tag{7.2}$$

In the absence of singles,

$$\hat{T}^{(1)} = \hat{T}_2^{(1)} \tag{7.3}$$

$$\hat{T}^{(2)} = \hat{T}_2^{(2)} + \hat{T}_3^{(2)} \tag{7.4}$$

for example, where $\hat{T}_4^{(2)}$ has been omitted since, although quadruples enter the second-order wavefunction, they may be factorized exactly into products of double excitations. For a more thorough introduction to perturbative analyses of CC theory, the reader is directed to one of the excellent reviews on the subject, such as Ref. [79].

The Traditional Coupled Cluster method may therefore be quite simply decomposed through orders of perturbation theory by constructing the n^{th}-order similarity-transformed Schrödinger Equation,

$$(e^{-\hat{T}}\hat{H}e^{\hat{T}})^{(n)}|\Phi_0\rangle = (\hat{H}e^{\hat{T}})_c^{(n)}|\Phi_0\rangle = E^{(n)}|\Phi_0\rangle. \tag{7.5}$$

Then, just as in regular TCC theory, the n^{th}-order energy is,

$$E_{\text{TCC}}^{(n)} = \langle \Phi_0 | (\hat{H} e^{\hat{T}})_c^{(n)} | \Phi_0 \rangle, \tag{7.6}$$

and the equations for the n^{th}-order amplitudes are also determined by projection onto the manifold of the appropriate excited determinants. In this way, the zeroth-order TCC energy is simply,

$$E_{\text{TCC}}^{(0)} = \langle \Phi_0 | (\hat{H} e^{\hat{T}})_c^{(0)} | \Phi_0 \rangle = \langle (\hat{F})_c \rangle = \langle \hat{F} \rangle, \tag{7.7}$$

since the cluster operator is at least first-order in the cluster amplitudes, and the first-order TCC energy is similarly,

$$\begin{aligned} E_{\text{TCC}}^{(1)} &= \langle \Phi_0 | (\hat{H} e^{\hat{T}})_c^{(1)} | \Phi_0 \rangle \\ &= \langle (\hat{V})_c \rangle + \langle (\hat{F} \hat{T}^{(1)})_c \rangle \\ &= \langle \hat{V} \rangle + \langle \hat{F} \hat{T}_2^{(1)} \rangle \\ &= \langle \hat{V} \rangle. \end{aligned} \tag{7.8}$$

By employing the same Møller-Plesset partitioning of the Hamiltionian as in the TCC case, it is similarly possible to decompose the VCC functional in orders of perturbation theory. The n^{th}-order VCC energy is simply,

$$E_{\text{VCC}}^{(n)} = \langle (e^{\hat{T}^\dagger} \hat{H} e^{\hat{T}})^{(n)} \rangle_L, \tag{7.9}$$

and the equations to be solved for the amplitudes are obtained by minimization of these energy expressions, rather than by projection. Since the cluster operator is at least first-order, the zeroth-order VCC energy is,

$$E_{\text{VCC}}^{(0)} = \langle \hat{F} \rangle_L = \langle \hat{F} \rangle. \tag{7.10}$$

The first-order VCC energy is,

$$\begin{aligned} E_{\text{VCC}}^{(1)} &= \langle \hat{V} \rangle_L + 2 \langle \hat{F} \hat{T}^{(1)} \rangle_L \\ &= \langle \hat{V} \rangle + 2 \langle \hat{F} \hat{T}^{(1)} \rangle \\ &= \langle \hat{V} \rangle + 2 \langle \hat{F} \hat{T}_2^{(1)} \rangle \\ &= \langle \hat{V} \rangle, \end{aligned} \tag{7.11}$$

where the second term vanishes by the Slater rules, and singles are again being ignored. The sum of the zeroth- and first-order energies therefore remains equal to the Hartree-Fock reference energy,

$$E_{\text{ref}} = \langle \hat{H} \rangle = \langle \hat{F} \rangle + \langle \hat{V} \rangle, \tag{7.12}$$

and, as is also the case for TCC, the first contribution to the correlation energy enters at second-order.

A comparison of the second-order TCC and VCC terms is now in order, and a result from standard Møller-Plesset theory must first be established. The equation satisfied by the first-order wavefunction,

$$\hat{V}|\Phi_0\rangle + \hat{F}|\Phi^{(1)}\rangle = E^{(1)}|\Phi_0\rangle + E^{(0)}|\Phi^{(1)}\rangle, \tag{7.13}$$

can be rearranged as,

$$(\hat{V} - E^{(1)}|\Phi_0\rangle + (\hat{F} - E^{(0)})|\Phi^{(1)}\rangle = 0, \tag{7.14}$$

or

$$(\hat{V} - \langle \hat{V} \rangle)|\Phi_0\rangle + (\hat{F} - \langle \hat{F} \rangle)|\Phi^{(1)}\rangle = 0, \tag{7.15}$$

It is convenient to introduce the normal-ordered Hamiltonian, which for the purposes of this discussion, is simply

$$\hat{H}_N = \hat{H} - \langle \hat{H} \rangle. \tag{7.16}$$

This is an operator for the *correlation* energy, since $\langle \hat{H}_N \rangle = 0$, and the $\hat{F}$ and $\hat{V}$ contributions may be similarly rewritten through,

$$\hat{F}_N = \hat{F} - \langle \hat{F} \rangle \tag{7.17}$$

$$\hat{V}_N = \hat{V} - \langle \hat{V} \rangle. \tag{7.18}$$

With these definitions, the equation satisfied by the first-order wavefunction becomes,

$$\hat{V}_N|\Phi_0\rangle + \hat{F}_N|\Phi^{(1)}\rangle = 0, \tag{7.19}$$

and, writing $|\Phi^{(1)}\rangle = \hat{T}^{(1)}|\Phi_0\rangle$, yields,

$$(\hat{V}_N + \hat{F}_N\hat{T}^{(1)})|\Phi_0\rangle = 0. \tag{7.20}$$

It has already been shown that the contributions to the reference energy are already fully accounted for by the zeroth- and first-order energies in both TCC and VCC, and the higher contributions are thus contributions to the correlation energy only. Normal-ordered operators may therefore be used without detriment in the remaining analysis.

From the previous discussions, the second-order TCC energy may be given as follows,

$$\begin{aligned} E^{(2)}_{\mathrm{TCC}} &= \langle(\hat{H}_N e^{\hat{T}})^{(2)}_c\rangle \\ &= \langle(\hat{F}\hat{T}^{(2)})_c\rangle + \langle(\hat{V}\hat{T}^{(1)})_c\rangle + \frac{1}{2!}\langle(\hat{F}(\hat{T}^{(1)})^2)_c\rangle \\ &= \langle(\hat{V}\hat{T}^{(1)}_2)_c\rangle \\ &= \langle\hat{V}\hat{T}^{(1)}_2\rangle, \end{aligned} \tag{7.21}$$

where, for no singles, the first and third terms vanish due to the Slater rules. The equation satisfied by $\hat{T}^{(1)}_2$ can be trivially determined by projection to be,

$$\langle\Phi^{ab}_{ij}|(\hat{H}_N e^{\hat{T}})^{(1)}_c|\Phi_0\rangle = \langle\Phi^{ab}_{ij}|(\hat{V}_N)_c|\Phi_0\rangle + \langle\Phi^{ab}_{ij}|(\hat{F}_N\hat{T}^{(1)}_2)_c|\Phi_0\rangle = 0, \tag{7.22}$$

or,

$$\langle\Phi^{ab}_{ij}|\hat{V}_N|\Phi_0\rangle + \langle\Phi^{ab}_{ij}|\hat{F}_N\hat{T}^{(1)}_2|\Phi_0\rangle = 0 \tag{7.23}$$

The second-order VCC energy, however, appears to take a different form,

$$\begin{aligned} E^{(2)}_{\mathrm{VCC}} &= 2\langle\hat{V}_N\hat{T}^{(1)}_2\rangle_L + \langle(\hat{T}^{(1)}_2)^\dagger\hat{F}_N\hat{T}^{(1)}_2\rangle_L \\ &= 2\langle\hat{V}_N\hat{T}^{(1)}_2\rangle + \langle(\hat{T}^{(1)}_2)^\dagger\hat{F}_N\hat{T}^{(1)}_2\rangle, \end{aligned} \tag{7.24}$$

where the unlinked diagram in the second term vanishes since $\langle\hat{F}_N\rangle = 0$. However, the criterion for the minimization of this functional, obtained by differentiation and setting the result equal to zero,

$$dE^{(2)}_{\mathrm{VCC}} = 2\langle\hat{V}_N d\hat{T}^{(1)}_2\rangle + 2\langle(\hat{T}^{(1)}_2)^\dagger\hat{F}_N d\hat{T}^{(1)}_2\rangle = 0, \tag{7.25}$$

reveals that $\hat{T}_2^{(1)}$ must satisfy the same equation as in TCC,

$$\langle \Phi_{ij}^{ab} | \hat{V}_N | \Phi_0 \rangle + \langle \Phi_{ij}^{ab} | \hat{F}_N \hat{T}_2^{(1)} | \Phi_0 \rangle = 0. \tag{7.26}$$

Furthermore, by using the result outlined above, or alternatively by manipulating the constituent diagrams with the amplitude equation, the VCC second-order energy can be shown to be, in fact, exactly equal to the TCC second-order expression due to the mutual cancellation of some of the VCC terms,

$$\begin{aligned} E_{\mathrm{VCC}}^{(2)} &= 2\langle \hat{V}_N \hat{T}_2^{(1)} \rangle + \langle (\hat{T}_2^{(1)})^\dagger \hat{F}_N \hat{T}_2^{(1)} \rangle \\ &= 2\langle \hat{V}_N \hat{T}_2^{(1)} \rangle - \langle (\hat{T}_2^{(1)})^\dagger \hat{V}_N \rangle \\ &= 2\langle \hat{V}_N \hat{T}_2^{(1)} \rangle - \langle \hat{V}_N \hat{T}_2^{(1)} \rangle \\ &= \langle \hat{V}_N \hat{T}_2^{(1)} \rangle \end{aligned} \tag{7.27}$$

It may be further established that both the TCC and VCC second-order energies are themselves equal to the second-order Møller-Plesset energy[79], such that both TCC and VCC are exactly correct to second-order in perturbation theory.

In fact, it is already well known[13, 54] that TCC and VCC agree not just to second-order, but to fourth-order in Møller-Plesset theory, and are fully correct to third-order, each omitting only the terms containing triple excitations from fourth-order. The purpose of the above discussion has been to illustrate how the mutual cancellation between the VCC contributions to the energy can yield the TCC contributions, and is particularly aesthetic because it brings to mind the cancellation between VCC diagrams for limiting systems that has been noted throughout this thesis. In addition, it is similarly known that not only is TCC equivalent to VCC to fourth-order, but at truncated levels, CCSD is equivalent to VCCSD, and CCSDT is equivalent to VCCSDT to this order. This has an *extremely* important practical consequence; the same minimalistic correction for the omission of fourth-order triples by CCSD, [T], is equally valid for VCCSD, such that both CCSD[T] and VCCSD[T] are both exactly correct to fourth-order in Møller-Plesset theory, as is intended by this correction. Therefore, if it can be additionally shown that QVCCD correctly generates all of the fourth-order VCCD contributions, the [T] correction becomes additionally valid for this Approximate VCC theory. This is the subject of the following section.

7.2 Perturbative Corrections for Approximate VCC Theories

At the end of the previous section, it was noted that the [T] correction for triple excitations is applicable not just to CCSD, but also to VCCSD, and becomes valid also for approximate VCC theories such as Quasi-Variational Coupled Cluster if QVCCD correctly agrees with VCCD to fourth-order in Møller-Plesset perturbation theory. This section demonstrates this to be true; the QVCCD method (and also the earlier AVCCD method) has been constructed specifically such that it agrees with VCCD to $\mathcal{O}(T^3)$, which has been demonstrated thoroughly in earlier chapters, and also such that it correctly accounts for the 1-electron $\mathcal{O}(T^4)$ terms. Although this last statement will not be proved in detail, due to to the large number of contributing diagrams, its validity will be made plausible in this section.

Previously the VCC energy was decomposed through second-order of perturbation theory. Consider now the contributions to specifically the VCCD energy through fourth-order of perturbation theory. The leading order contribution to the VCCD energy is the reference energy,

$$E^{(0)}_{\mathrm{VCCD}} + E^{(1)}_{\mathrm{VCCD}} = \langle \hat{F} \rangle + \langle \hat{V} \rangle = \langle \hat{H} \rangle. \tag{7.28}$$

The reference energy is the first term present in the QVCCD functional. Therefore QVCCD is correct and equivalent to VCCD to at least first-order. Similarly, it was previously shown that,

$$E^{(2)}_{\mathrm{VCCD}} = 2\langle \hat{V}_N \hat{T}^{(1)}_2 \rangle + \langle (\hat{T}^{(1)}_2)^\dagger \hat{F}_N \hat{T}^{(1)}_2 \rangle = 2\langle \hat{V} \hat{T}^{(1)}_2 \rangle + \langle (\hat{T}^{(1)}_2)^\dagger (\hat{F} - \langle \hat{F} \rangle) \hat{T}^{(1)}_2 \rangle. \tag{7.29}$$

Both of these terms are contained in the CEPA(0) functional, which is itself contained by QVCCD. Hence, QVCCD is also equivalent to VCCD to at least second-order. The third-order VCCD energy,

$$E^{(3)}_{\mathrm{VCCD}} = 2\langle \hat{V}_N \hat{T}^{(2)}_2 \rangle + \langle (\hat{T}^{(1)}_2)^\dagger \hat{V}_N \hat{T}^{(1)}_2 \rangle_L + 2\langle (\hat{T}^{(1)}_2)^\dagger \hat{F}_N \hat{T}^{(2)}_2 \rangle_L \tag{7.30}$$

is similarly captured by CEPA(0), and hence also by QVCCD.

Therefore, only an examination of the fourth-order VCCD terms remains. Of these terms, several are captured by CEPA(0),

$$E^{(4)}_{\mathrm{VCCD}} \longleftarrow 2\langle(\hat{T}_2^{(1)})^\dagger \hat{F}_N \hat{T}_2^{(3)}\rangle_L + \langle(\hat{T}_2^{(2)})^\dagger \hat{F}_N \hat{T}_2^{(2)}\rangle_L + 2\langle \hat{V}_N \hat{T}_2^{(3)}\rangle + 2\langle(\hat{T}_2^{(1)})^\dagger \hat{V}_N \hat{T}_2^{(2)}\rangle_L \tag{7.31}$$

but the remaining terms,

$$E^{(4)}_{\mathrm{VCCD}} \longleftarrow \langle(\hat{T}_2^{(1)})^\dagger \hat{V}_N (\hat{T}_2^{(1)})^2\rangle + \left(\frac{1}{2!}\right)^2 \langle((\hat{T}_2^{(1)})^\dagger)^2 \hat{F}_N (\hat{T}_2^{(1)})^2\rangle_L \tag{7.32}$$

are not. Those terms captured by CEPA(0) are automatically present in QVCCD, but the remaining terms require further discussion. The first of these terms enters from the $\mathcal{O}(T^3)$ VCCD term, $\langle \hat{T}_2^\dagger \hat{H}_N \hat{T}_2^2\rangle$, which is generated exactly by QVCCD, and also AVCCD, although not LPFD. The other term arises from the 1-electron $\mathcal{O}(T^4)$ terms that have been mentioned several times. In fact, it is a consequence of the fact that ${}_2\hat{T}$ generates the correct $\mathcal{O}(T^3)$ VCCD terms that ${}_1\hat{T}$ generates the correct 1-electron $\mathcal{O}(T^4)$ terms. While some of the 2-electron $\mathcal{O}(T^4)$ terms are omitted by QVCCD, these terms are at least fifth-order in perturbation theory and are also omitted by CCD, and it is impossible to include them without violating the $\mathcal{O}(o^2v^4)$ complexity criterion.

In order to further justify the point that QVCCD correctly generates the 1-electron $\mathcal{O}(T^4)$ terms, which it is necessary to get right if QVCCD is to be equivalent to VCCD to fourth-order in perturbation theory, consider one of the $\mathcal{O}(T^4)$ terms generated by the $\mathcal{C}$ transformation in $\langle {}_1\hat{T}_2^\dagger \hat{F}_N {}_1\hat{T}_2\rangle$.

$$\begin{aligned}
\langle {}_1\hat{T}_2^\dagger \hat{F}_N {}_1\hat{T}_2\rangle &= \langle {}_1\hat{T}_2^\dagger \hat{F} {}_1\hat{T}_2\rangle_L \\
&\leftarrow -\frac{1}{2} {}_1T^{ij}_{ab} {}_1T^{ab}_{ik} f^k_j \\
&\leftarrow -\frac{1}{4}(c\eta\mathbf{T})^{ij}_{ab} T^{ab}_{ik} f^k_j - \frac{1}{4} T^{ij}_{ab} (c\eta\mathbf{T})^{ab}_{ik} f^k_j \\
&= -\frac{1}{8} c\eta^{ij}_{mn} T^{mn}_{ab} T^{ab}_{ik} f^k_j - \frac{1}{8} T^{ij}_{ab} c\eta^{mn}_{ik} T^{ab}_{mn} f^k_j \\
&= -\frac{1}{4} c\eta^{ij}_{mn} c\eta^{mn}_{ik} f^k_j - \frac{1}{4} c\eta^{ij}_{mn} c\eta^{mn}_{ik} f^k_j \\
&= -\frac{1}{2} c\eta^{ij}_{mn} c\eta^{mn}_{ik} f^k_j
\end{aligned} \tag{7.33}$$

The coefficient of $-1/2$ (or $-1/8$ upon inserting the definition of each ${}_c\eta$), agrees exactly with the weighting of the generated term as it appears in VCCD, as

can be confirmed diagrammatically. All the 1-electron $\mathcal{O}(T^4)$ terms are similarly captured by QVCCD, and with the correct coefficients, and QVCCD (and also AVCCD but not LPFD) is therefore equivalent to VCCD to fourth-order in Møller-Plesset perturbation theory.

With this result, QVCCD inherits the validity of the [T] correction from VCCD, which was itself inherited from CCD. More precisely, both OQVCCD and BQVCCD may make use of the same [T] correction as BCCD, with the effect that OQVCCD[T] and BQVCCD[T], in addition to their numerous other advantageous methodological properties, are also fully correct to fourth-order in Møller-Plesset perturbation theory, the same as CCSD[T] or BCCD[T]. It is also noteworthy that there is no difference between the [T] correction and the (T) correction when the single excitations vanish, which, of course, occurs for the OQVCCD scheme, since they differ only in terms that contain single excitations. This defines the Optimized-orbital Quasi-Variational Coupled Cluster Doubles with Perturbative Triples (OQVCCD(T)) method, which possesses the following, extremely theoretically attractive set of methodological properties.

- OQVCCD(T) is a unique theory, containing no arbitrary parameters.
- The iterative step in the calculation of the ground-state energy is performed by variational minimization of a functional, such that the error is second-order in any remaining errors in the cluster amplitude parameters.
- The functional and perturbative correction contain fully linked terms only, such that calculated energies are rigorously extensive.
- It is exactly equivalent to FCI for 2 electrons or 2 holes, systems for which the cluster operator $\hat{T} = \hat{T}_1 + \hat{T}_2$ is complete.
- The functional yields a scalar energy that is invariant to rotations in the underlying orbital spaces $\{\psi_i\}$ and $\{\psi_a\}$.
- The energy is not a true upper bound on the exact ground-state Schrödinger energy eigenvalue, but variational minimization of the functional gives a theory that satisfies the generalized Hellmann-Feynman theorem[67].

- The doubles-only theory is equivalent to VCCD to $\mathcal{O}(T^3)$, and, as a consequence, the one-electron $\mathcal{O}(T^4)$ terms are also constructed correctly.
- Furthermore, by construction, balanced VCCD contributions are generated through all orders of the cluster amplitudes, such that the functional potentially inherits approximate upper bound character from its parent VCCD method.
- It is correct to fourth-order in Møller-Plesset perturbation theory[3].
- The limiting computational complexity of the method is iterative $\mathcal{O}(o^2v^4)$ and non-iterative $\mathcal{O}(o^3v^4)$ with $\mathcal{O}(o^2v^2)$ storage, the same as CCSD(T).

The OQVCCD and OQVCCD(T) methods will be thoroughly benchmarked against their competitors in the section that follows.

7.3 Benchmark Results

As with the other data presented in this thesis, most calculations in this chapter have been performed with the Molpro[86, 87] quantum chemistry software package. For each molecule studied, one-dimensional cuts of the potential energy surface obtained with various single-reference coupled-cluster methodologies are compared with those obtained from internally-contracted multireference configuration interaction[29, 30] (MRCI) calculations. These reference calculations use complete active space reference wavefunctions where the active space consists of the atomic valence orbitals, and the energy is corrected using the approximate extensivity correction of Davidson[113] (MRCI+Q). Using the GAMESS[114] package, the CR-CC(2,3) method[112] is also examined. CEPA results have additionally been obtained using Orca[115].

To begin, two simple single bond breaking examples are investigated. Calculated potential energy curves for BH with the cc-pVQZ basis are given in Figure 7.1 and for HF with the aug-cc-pVQZ basis in Figure 7.2. In both examples, the CCSD(T) method becomes poor as the bond is stretched. This is a well-known problem associated with the (T) correction, since it becomes singular when

the highest occupied and lowest unoccupied molecular orbitals become degenerate. However, while CCSD and OQVCCD perform similarly, the OQVCCD(T) method fares significantly better than CCSD(T). While this is promising, comparison with BCCD and BCCD(T) shows that this behaviour may be a result of the different orbitals in use; CCSD(T) uses Hartree-Fock orbitals, whereas BCCD(T) uses Brueckner orbitals and OQVCCD(T) uses variationally optimal orbitals. Equivalently, these differences are a result of the different treatment of single excitations, due to the Thouless theorem[99], which states that any two single-determinantal wavefunctions, $|\Phi\rangle$ and $|\Phi'\rangle$, may be related by $|\Phi'\rangle = e^{\hat{T}_1}|\Phi\rangle$. This behaviour has been noted previously, for example by Nooijen and Le Roy, who found that the use of Brueckner orbitals improved the triples corrections substantially in HF, BeO, CN and BN[18]. In the calculations that follow, the BCCD and BCCD(T) methods are used to identify those systems for which the choice of orbitals affects the triples corrections more than the differences in the doubles-only theories. It should be noted, however, that the Brueckner orbitals are not always close to the variationally optimal orbitals[100], and it has been checked, for additional clarity, that Brueckner Quasi-Variational Coupled Cluster Doubles (BQVCCD) performs in agreement with OQVCCD in the examples that follow.

Next, the spectroscopic constants for a selection of diatomic molecules, are examined, and presented in Table 7.1. The CCSD and OQVCCD results are of similar quality, as are the CCSD(T) and OQVCCD(T) results. This is to be expected due to the similarity of the potential energy curves around equilibria, evident in later figures. It is noteworthy that comparing the spectroscopic constants obtained from the BQVCCD, BCCD, BQVCCD(T) and BCCD(T) methods reveals deficiencies that arise in the current QVCCD approximation to VCCD, which manifest as incomplete recovery of dynamic correlation energy relative to CCD, leading to a slightly poorer description of the equilibrium region of potential energy surfaces. These defects are, however, quite small.

A first example involving strong non-dynamic correlation is the case of the symmetric stretching of a double bond, for which $H_2Si = SiH_2$ constrained to a planar D_{2h} geometry is taken. The Si-H bond length and bond angle are optimized

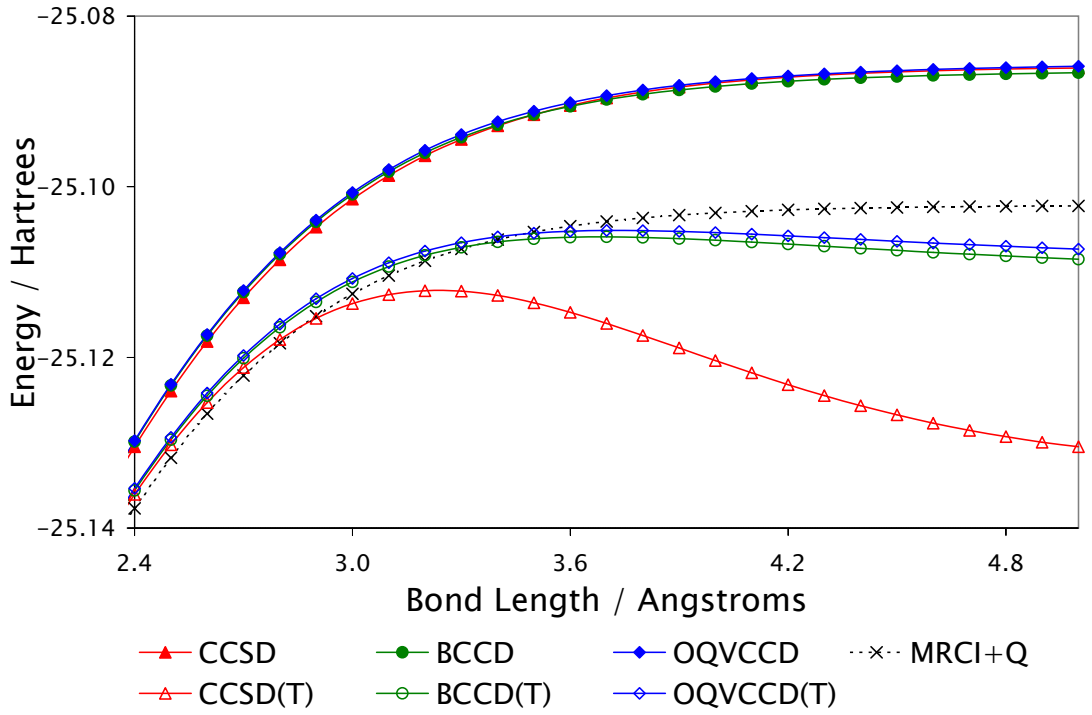


Figure 7.1: Calculated potential energy curves for BH with the cc-pVQZ basis set.

at the CCSD level of theory for each Si-Si distance, and the results of the various methods are illustrated in Figure 7.3.

The first noteworthy feature of this graph is the strong divergence between OQVCCD and the TCC methods, CCSD and BCCD, that occurs from around 3.6Å. Naïvely, it may appear that CCSD and BCCD are closer to MRCI+Q, and therefore perform better than OQVCCD. However, the more slowly decreasing slope of the OQVCCD curve, in fact, mimics the divergence of VCCSD from FCI evident in previous examples, such as Figure 2.1, and that can be observed in other VCCSD minimal basis benchmarking exercises[95]. The difference in this example is that the TCC methods do not fail as severely. The divergence of CCSD from OQVCCD therefore reflects the divergence of CCSD from VCCSD. Further supporting evidence will follow in other examples.

Another point of particular interest, however, is that, correlated with the divergence of the TCC methods from OQVCCD (or VCCSD) around 3.6Å, is the onset of a catastrophic failure of the CCSD(T) and BCCD(T) methods to describe the regime of dissociation. Both triples-corrected TCC methods pre-

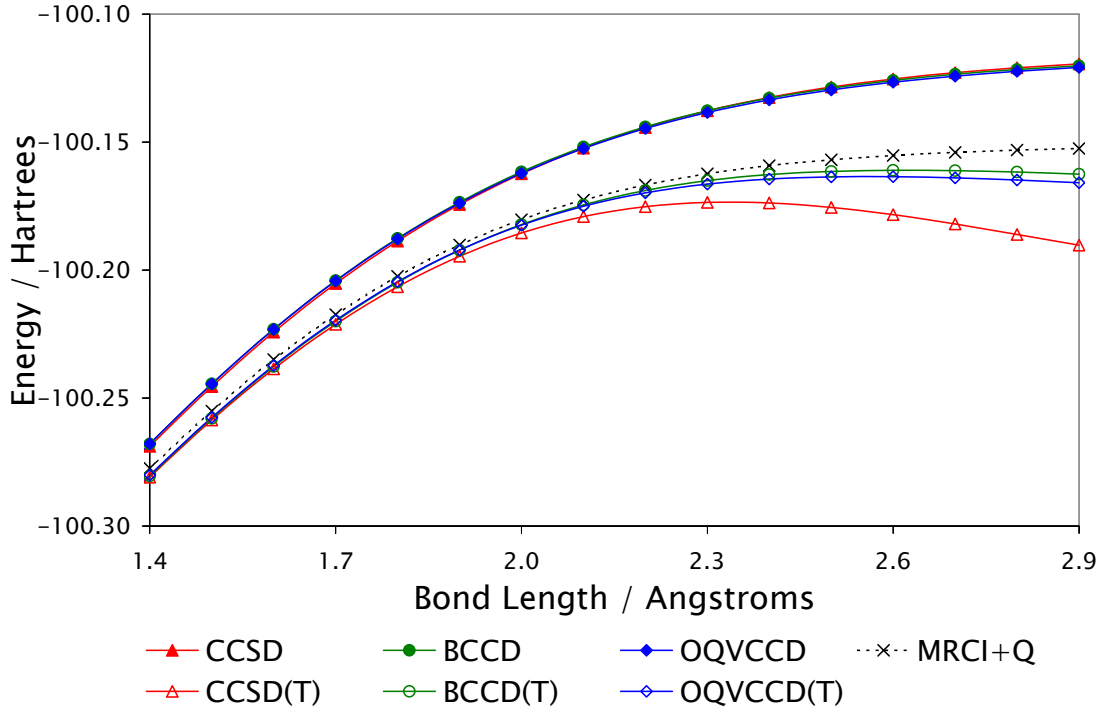


Figure 7.2: Calculated potential energy curves for HF with the aug-cc-pVQZ basis set.

dict unphysical maxima in the potential energy curve around 3.8Å, and subsequently diverge to energies significantly below MRCI+Q. Although insufficient CR-CC(2,3) points were able to be acquired in order to fully evaluate the performance of this (more advanced) method, which has yielded some impressive results for the breaking of single bonds[117], by 4.2Å, the CR-CC(2,3) curve has already crossed the MRCI+Q curve, and it is therefore reasonable to assume that it performs better than, but qualitatively similarly to CCSD(T). In contrast, the OQVCCD(T) method displays no significant problems at this, or even longer bond lengths, and instead continues in a physically correct manner, remaining just above the MRCI+Q curve throughout.

Another class of systems with which the TCC-based *ab initio* methods are known to struggle is the simultaneous breaking of several single bonds. One classic example is the symmetric stretching of the two O-H bonds in water. For the purposes of benchmarking, however, it is fitting to examine some more severe examples of non-dynamic correlation, such as the case of the symmetric stretching of all 3 of the N-H bonds in ammonia, NH_3. Calculated potential energy curves for this system are given in Figure 7.4.

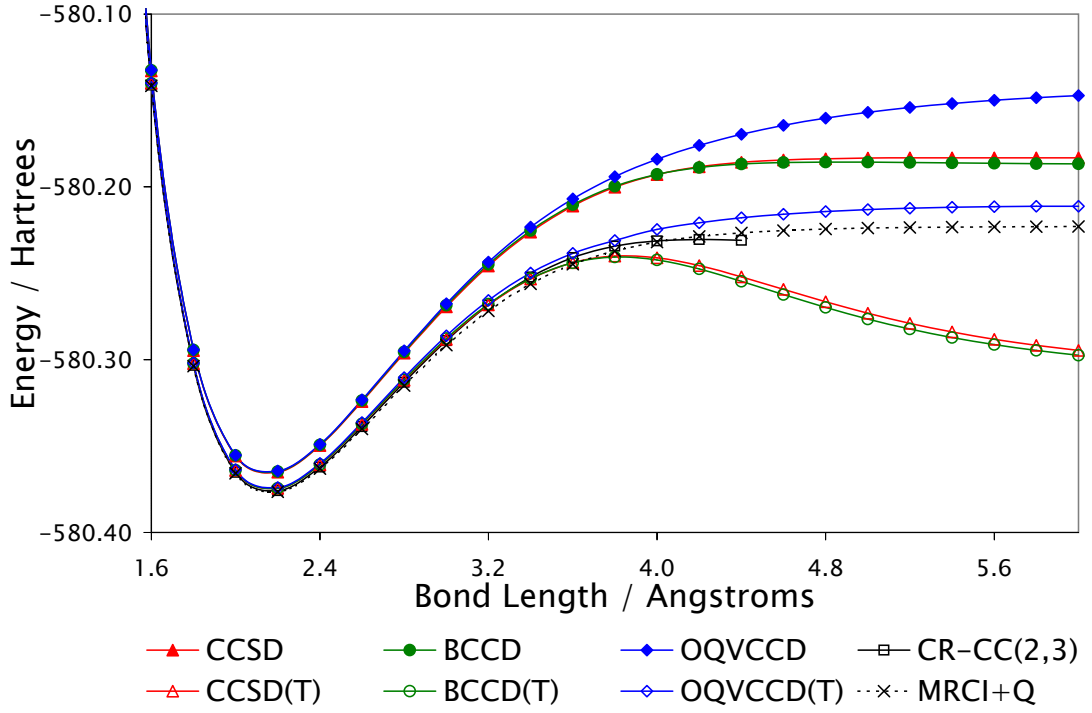


Figure 7.3: Calculated potential energy curves for D_{2h} $H_2Si = SiH_2$ as a function of the Si = Si bond length with the cc-pV(D+d)Z basis set.

The CCSD and BCCD methods both encounter difficulties as the bonds in NH_3 are stretched; the BCCD method undergoes a non-variational collapse to energies below the MRCI+Q curve, and the CCSD curve becomes unstable and begins to increase in energy too sharply from 2.2 Å. Both curves are clearly wrong in comparison to the MRCI+Q curve. However, the OQVCCD curve is smooth and continues fairly parallel to MRCI+Q throughout. These differences are magnified and made significantly more obvious when the (T) correction is added to each method; both CCSD(T) and BCCD(T) then diverge unphysically, whereas OQVCCD(T) is in excellent agreement with MRCI+Q, representing a significant improvement over the methods based on TCC. While CR-CC(2,3) does not demonstrate a failure as pronounced as in CCSD(T), it exhibits instabilities similar to the CCSD method on which it is based. The CEPA-2 method[11, 13] has additionally been examined for this example, and the results show it to diverge from MRCI+Q and become catastrophically unphysical at bond lengths even shorter than for which the breakdown of CCSD(T) occurs.

The symmetric stretching of the C-H bonds in ethene, in which four bonds are now simultaneously broken, represents an even more extreme test case, and

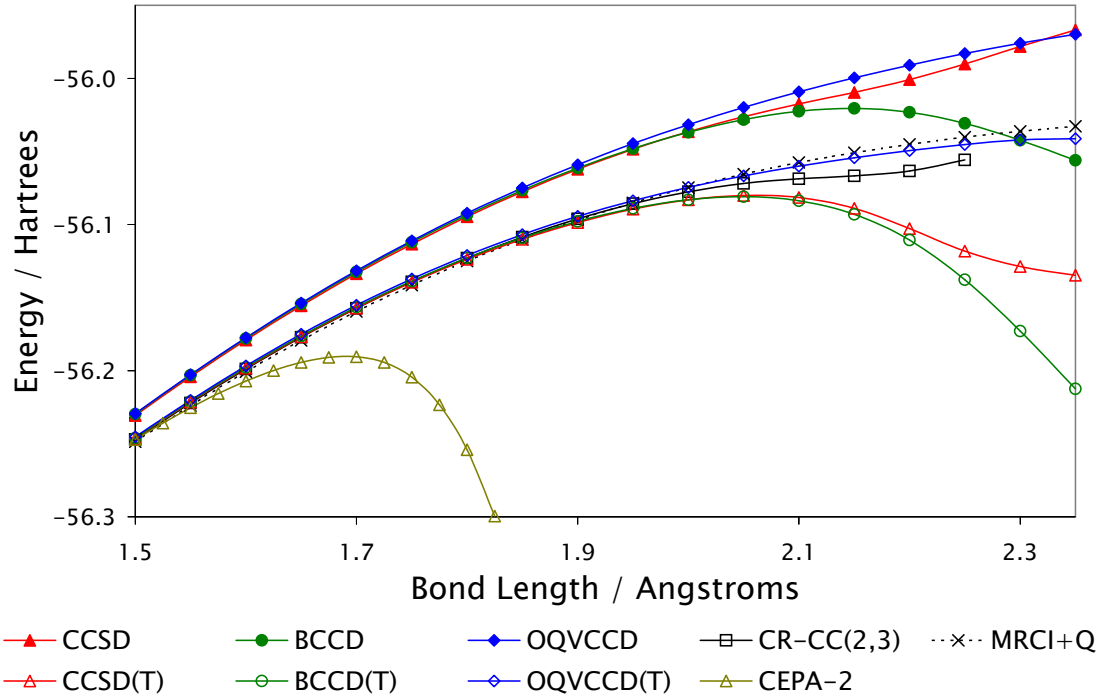


Figure 7.4: Calculated potential energy curves for the symmetric stretching of NH_3 with the cc-pVTZ basis set.

results are illustrated in Figure 7.5. All methods tested are capable of treating the equilibrium region of the potential energy curve to the same level of accuracy, but the CCSD and BCCD methods both become unphysical and qualitatively incorrect in comparison to MRCI+Q from 2.2 Å, even before the corrections for triples are added. In contrast, OQVCCD does not predict an unphysical maximum, and the curve continues as one would expect from physical intuition.

When the triples corrections are added, CCSD(T) and BCCD(T) both diverge even more rapidly. The CR-CC(2,3) performs slightly better, and is apparently capable of treating the dynamic correlation to the level of CCSD(T), but diverging at long bond lengths with slope more like that of CCSD. It is still unphysical, and therefore wrong, at these long bond lengths, however. The OQVCCD(T) method, however, appears to be just as stable and physically correct as OQVCCD, but also accounts for much of the remaining dynamic correlation, and therefore remains not just in qualitative agreement, but also in excellent quantitative agreement with MRCI+Q. The contrast between the TCC-based methods and OQVCCD(T) is therefore quite extraordinary in this example.

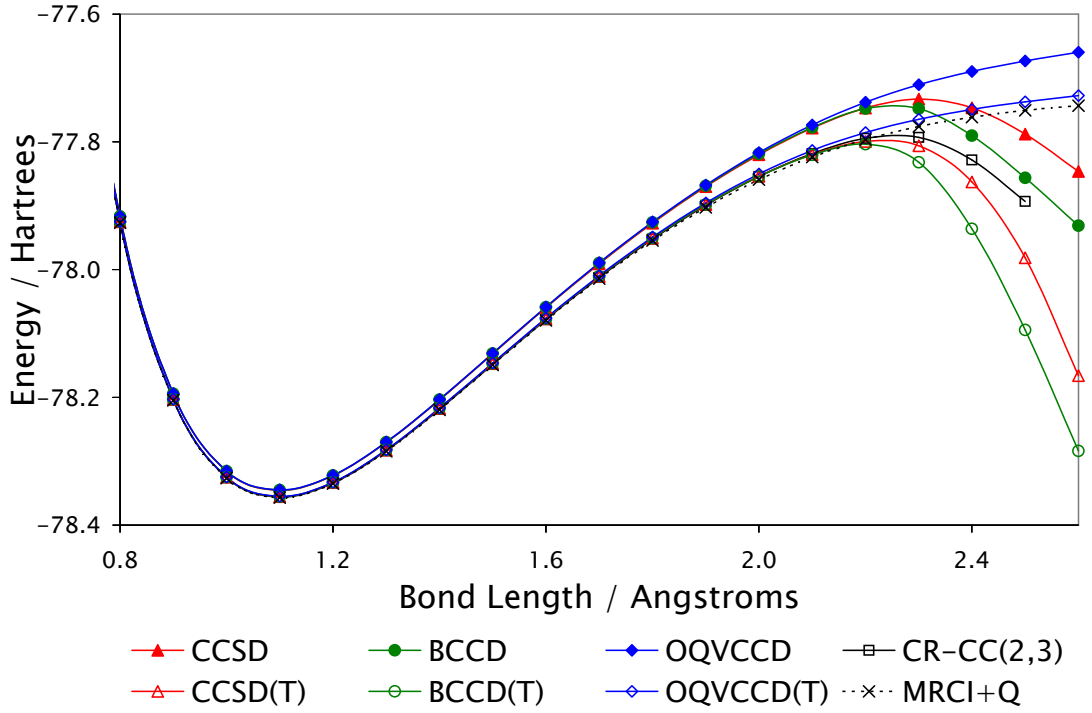


Figure 7.5: Calculated potential energy curves for the symmetric stretching of the C-H bonds in ethene with the cc-pVDZ basis set.

As a further example of simultaneous bond breaking, and of the extent to which OQVCCD approximates VCCSD well in non-dynamic situations, consider the stretching of the N-H and O-H bonds in NH_2OH, for which results are illustrated in Figure 7.6 with the minimal STO-3G basis, with which VCCSD results can again be obtained. The overall conclusions for this system are the same as for ammonia and ethene; OQVCCD predicts a physically correct potential energy curve, whereas CCSD and BCCD do not, and the (T) correction improves the OQVCCD potential energy curve, whereas its application to either CCSD or BCCD leads to a huge overestimation of the correlation energy. The effect of the (T) correction appears to be small, especially for OQVCCD, in this example, but this is, in fact, simply an artefact of the minimal basis set in use, and the (T) correction becomes more important as the basis set is made larger. In this basis, another feature becomes apparent, however; there is a shift towards higher energies due to the use of either orbital optimization or Brueckner orbitals, which motivates further development of explicit treatments of single excitations in the Approximate VCC family of methods.

In order to better understand the degree to which OQVCCD out-performs

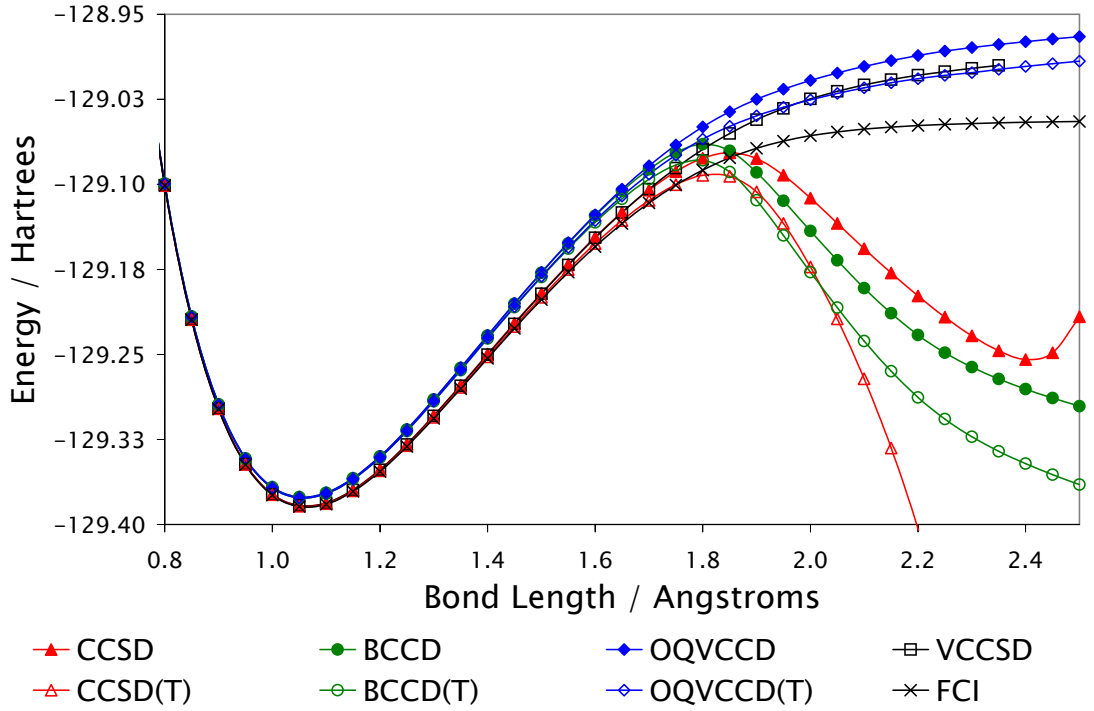


Figure 7.6: Calculated potential energy curves for the symmetric stretching of the N-H and O-H bonds in NH_2OH with the STO-3G basis set.

CCSD for the treatment of non-dynamic correlation, a more detailed study of the predicted potential energy curves of the lowest singlet state of C_2 is now presented, another difficult test of single-reference Post-Hartree-Fock methods. Calculated potential energy curves for this system are first given in Figure 7.7 with the aug-cc-pVQZ basis.

As in the case of Si_2H_4, the OQVCCD method diverges from the CCSD method at long bond lengths; unlike CCSD, the curve does not plateau around 2.4 Å, leading OQVCCD to predict a larger dissociation energy than CCSD. This, again, reflects the underlying divergence between CCSD and VCCSD that becomes apparent if the system is examined in a minimal basis[95]. When the triples corrections are added, CCSD(T) and OQVCCD(T) are similar around equilibrium, but CCSD(T) becomes unphysical from approximately 2.3 Å. On the other hand, OQVCCD(T) remains physically correct throughout, and in excellent agreement with MRCI+Q even to extremely long bond lengths. It is noteworthy that, just as in the previous examples, the CCSD(T) and BCCD(T) methods turn over precisely when CCSD and BCCD begin to diverge from OQVCCD, independent of whether the CCSD and BCCD methods themselves become unphysical.

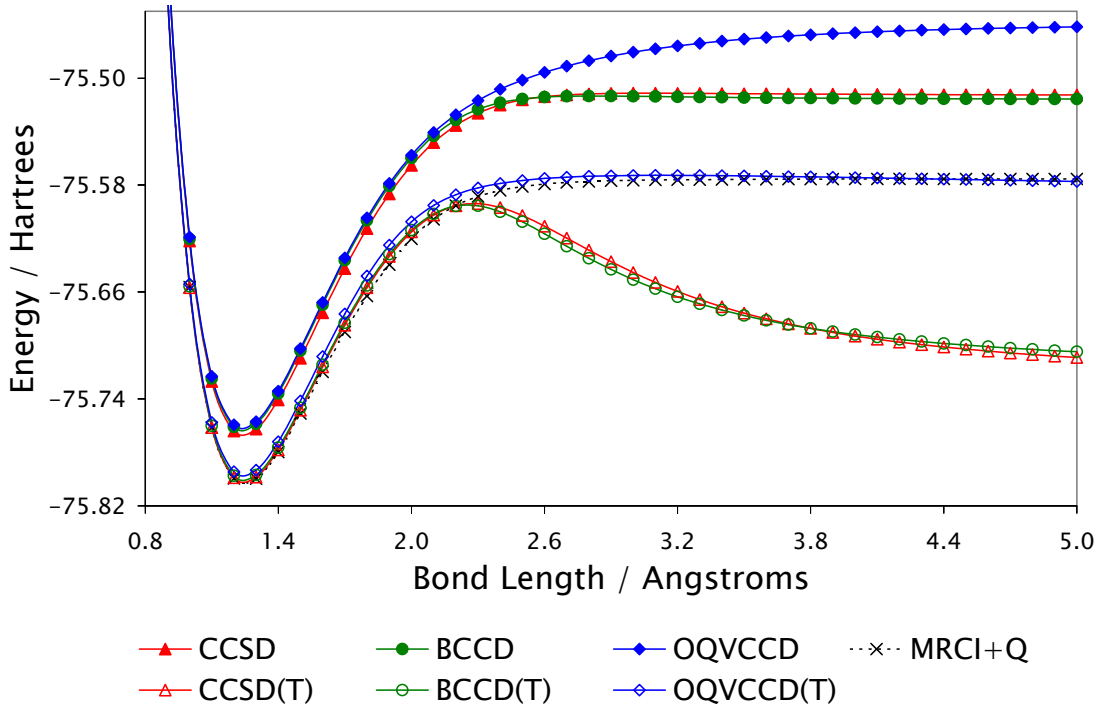


Figure 7.7: Calculated potential energy curves for C_2 with the aug-cc-pVQZ basis set.

To further support these results, a plot of the calculated forces is additionally presented in Figure 7.8, calculated by finite-difference differentiation, and shows the OQVCCD(T) curve to be almost flat, predicting a correct dissociation of the molecule, well beyond the region in which CCSD(T) and BCCD(T) become exceedingly inaccurate. Furthermore, a plot of the square norms of the cluster amplitudes for the methods CCSD, BCCD and OQVCCD is presented in Figure 7.9. The OQVCCD square norm increases slowly with decreasing gradient from around 2.4 Å, reaching a value of 2.40 at 1000 Å, suggesting that only small changes will occur at even longer bond lengths. In contrast, the CCSD and BCCD square norms approach their respective limits more slowly than OQVCCD, and this appears to be correlated with the failure of the CCSD(T) and BCCD(T) methods.

In Figure 7.10, calculated potential energy curves for N_2 with the aug-cc-pVQZ basis set are given. This system is a notorious test of single-reference correlation methods, and both CCSD and BCCD clearly provide inadequate descriptions of the breaking of the triple bond in this strongly-correlated example. When the (T) correction is added to either CCSD or BCCD, the problem is magnified. Even

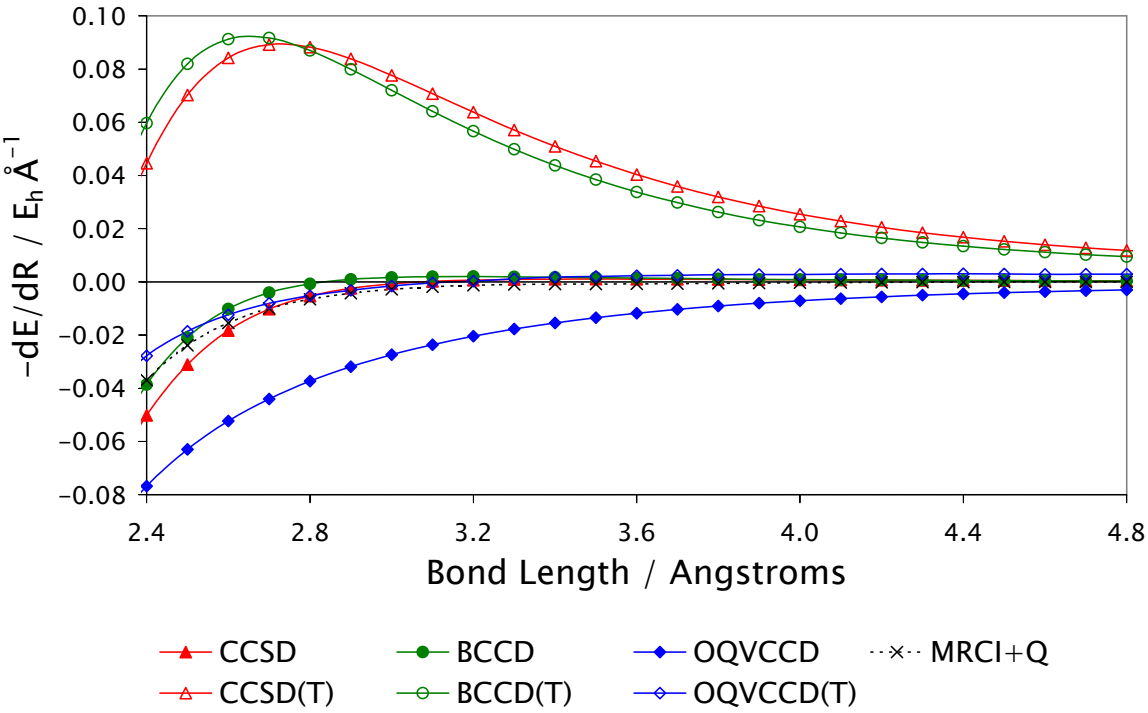


Figure 7.8: Calculated force curves for C_2 with the aug-cc-pVQZ basis set.

CR-CC(2,3) collapses in a non-variational fashion to energies significantly below the MRCI+Q curve. However, the OQVCCD curve, again, does not manifest the same unphysical maximum in the potential energy curve as the TCC methods, and OQVCCD(T) is in outstanding agreement with MRCI+Q out to internuclear separations where the triple covalent bond is essentially broken.

In Figure 7.11, results are presented with the smaller cc-pVDZ basis, in which it is possible to perform comparative CCSDT and CCSDTQ calculations using Kállay's MRCC program[118]. The results are presented as a plot of the errors relative to MRCI+Q. From this data, it is apparent that CCSDT also fails completely, but inclusion of the effects of connected quadruple excitations yields a qualitatively correct curve. The behaviour of OQVCCD(T) is simply exceptional, with this iterative $\mathcal{O}(N^6)$ plus non-iterative $\mathcal{O}(N^7)$-scaling theory outperforming iterative $\mathcal{O}(N^8)$ CCSDT, and rivalling iterative $\mathcal{O}(N^{10})$ CCSDTQ. Even though one might expect the perturbative part of OQVCCD(T) to become unreliable as the highest occupied and lowest unoccupied molecular orbitals approach degeneracy, the rate at which OQVCCD(T) diverges from MRCI+Q slows at long bond lengths, and the present results therefore challenge the conventional wisdom that

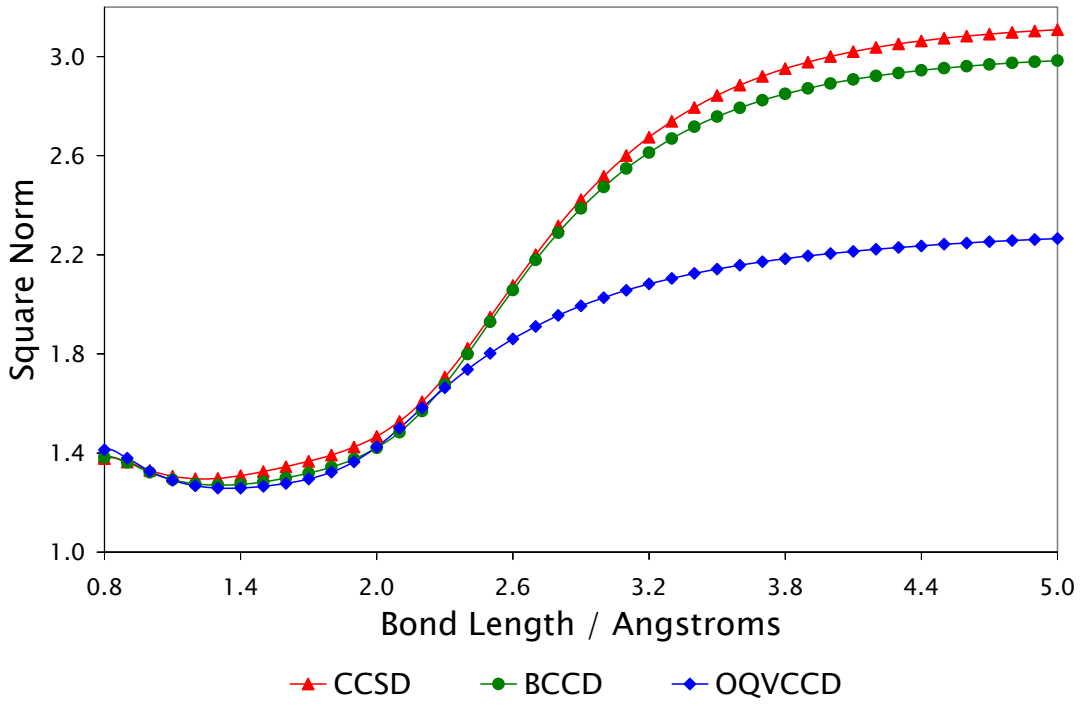


Figure 7.9: Amplitude square norms for C_2 with the aug-cc-pVQZ basis set.

a full, non-perturbative treatment of triple and quadruple excitations is required to correctly dissociate this system.

The acetylene molecule, C_2H_2, which has been used as a prototypical example of multiple bond breaking throughout this thesis, possesses an electronic structure that is analogous to the N_2 molecule, and thus represents a similarly extreme test of molecular electronic structure methods when the triple bond is stretched. The results of calculations on this system are shown in Figure 7.12, for a fixed C-H bond length of 1.06Å. From approximately 2.2 Å, none of the methods based on TCC are correct, each predicting an unphysical maximum in the potential energy curve followed by a non-variational collapse to energies below MRCI+Q. The effect of the (T) correction on each of these methods is to push the energy even lower, causing the problem to become magnified. The CEPA-2 method also diverges to unphysically low energies on stretching, and the onset of this behaviour occurs at even shorter bond lengths (from around 1.6 Å). In contrast, the OQVCCD method does not appear to degrade significantly in quality as the bond is stretched, predicting a potential energy curve with the characteristic VCCSD shape[95], and the additional (T) perturbative correction of the energy

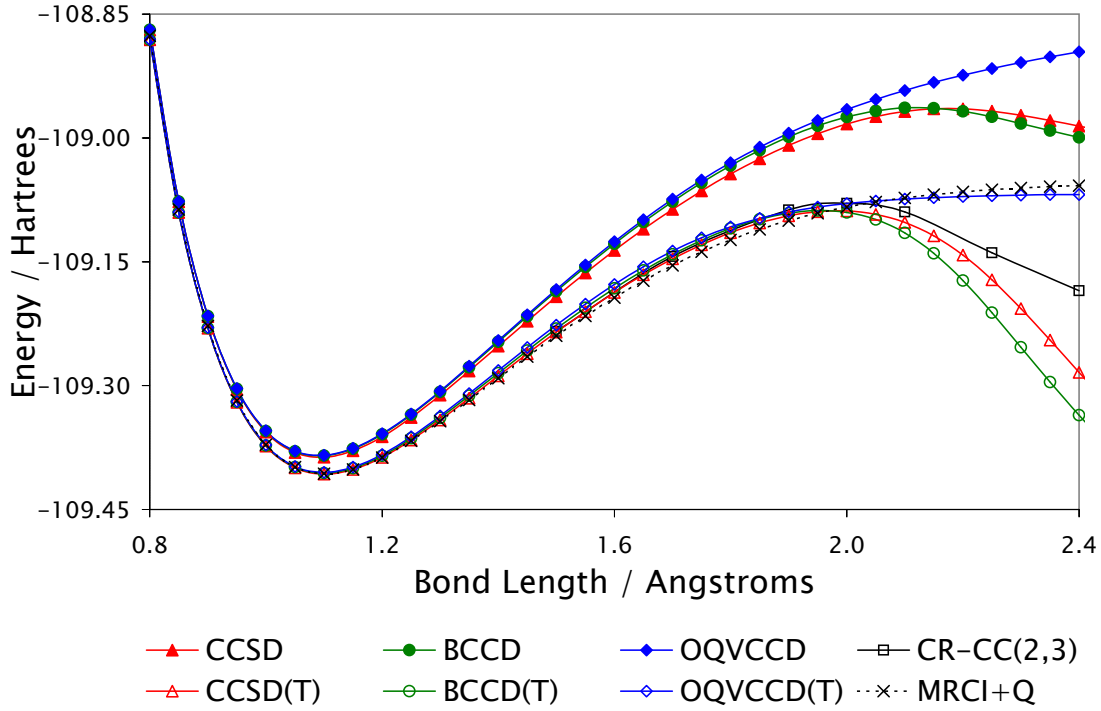


Figure 7.10: Calculated potential energy curves for N_2 with the aug-cc-pVQZ basis set.

results in a predicted curve that is in outstanding agreement with MRCI+Q.

As a final example of multiple bond breaking, consider the model system cyclo-H_8, in which eight hydrogen atoms are arranged at the vertices of a regular octagon, with edge length R. As R is increased, the 8 partial bonds break simultaneously. In addition, the physically correct behaviour in the dissociation limit is for one electron to localize on each of the atomic sites, and the restricted Hartree-Fock reference wavefunction used for each of the single-reference calculations therefore suffers from severe ionic contamination[59]. With a total of 48 determinants approximately equally important at dissociation, this model system represents possibly the most severe test of single-reference *ab initio* methods for non-dynamic electron correlation examined in this thesis.

This severity is quite apparent in the TCC potential energy curves, presented in Figure 7.13, with all of CCSD, CCSD(T), BCCD and BCCD(T) nonsensical from around 1.8Å. It is a testament to the robust nature of electronic structures that can be calculated with even an approximately-fulfilled variational upper bound property that OQVCCD should perform at all well in this system, since

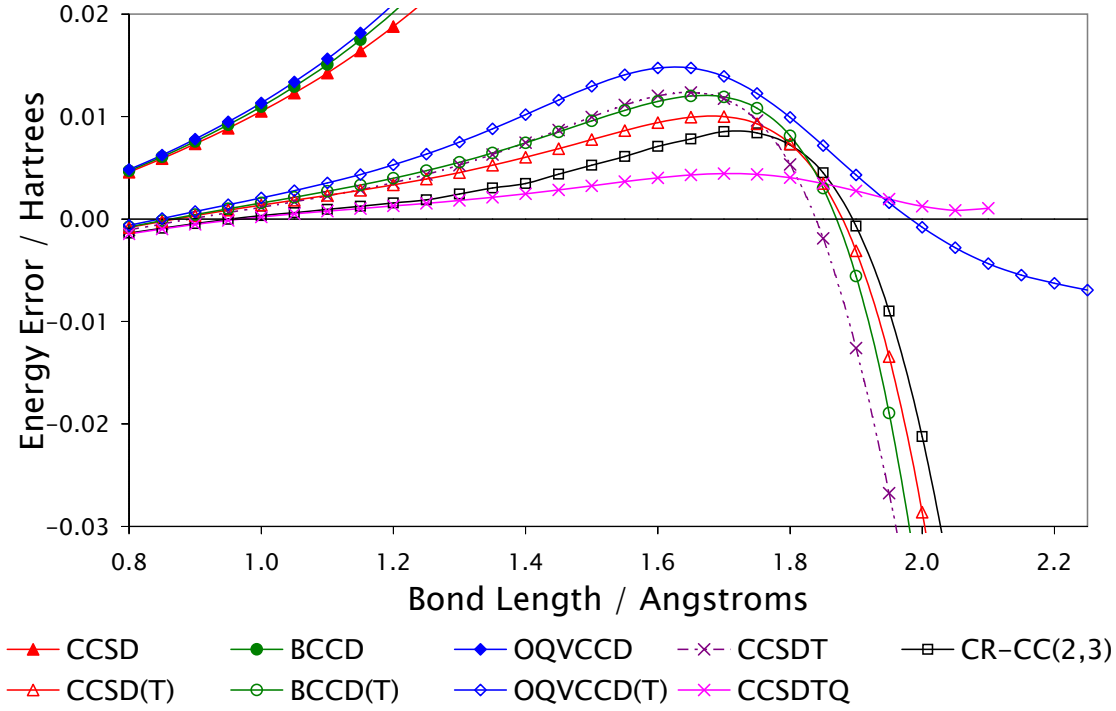


Figure 7.11: Errors relative to MRCI+Q for N_2 with the cc-pVDZ basis set.

it is, in fact, physically correct throughout. Dynamic correlation is weak in this system, since the electrons should be spatially local at long distances, and the effect of (T) is therefore small, but the agreement of OQVCCD(T) with MRCI+Q is simply magnificent.

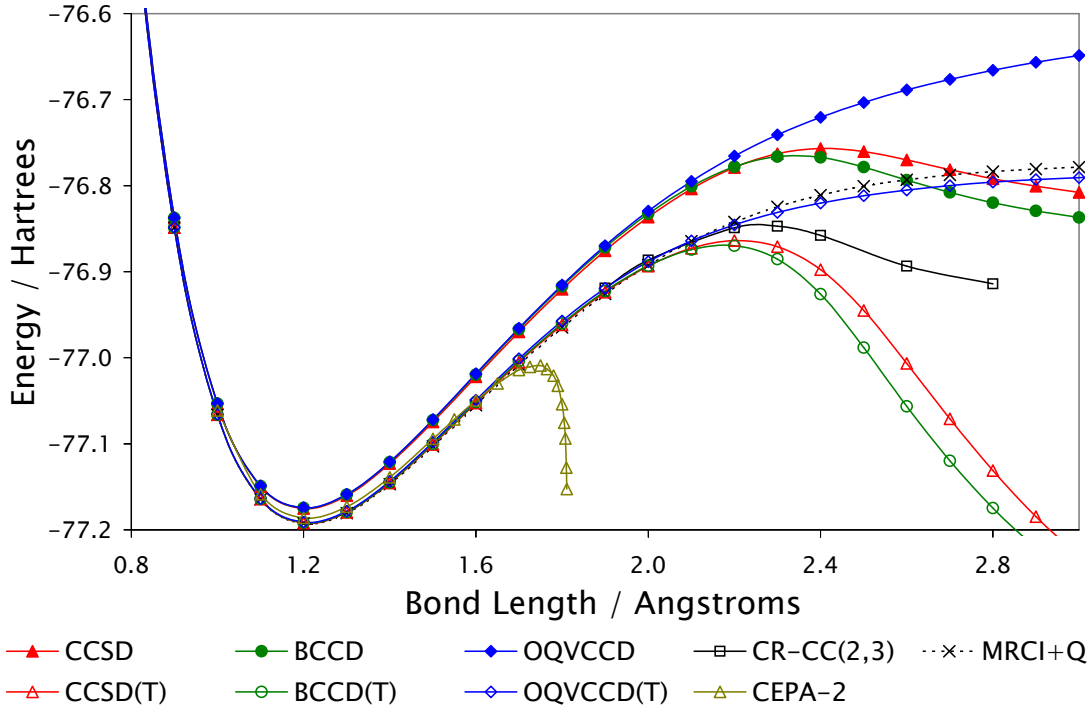


Figure 7.12: Calculated potential energy curves for the stretching of the carbon-carbon triple bond in acetylene, C_2H_2, with the C-H bond length fixed at 1.06Å, and with the aug-cc-pVTZ basis set.

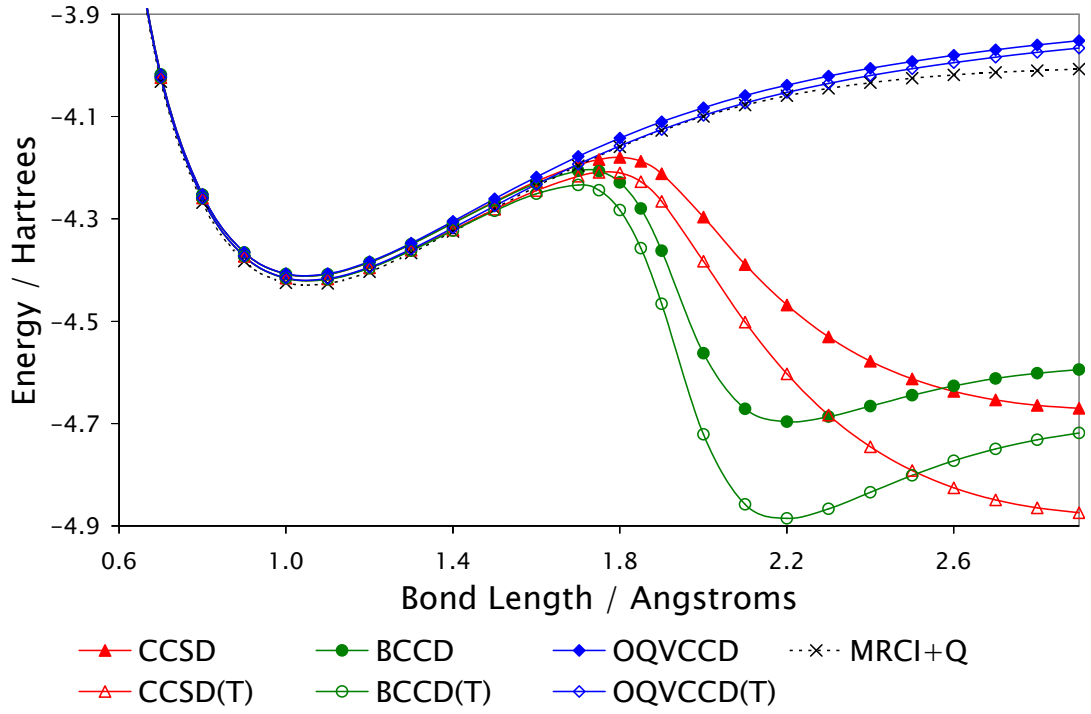


Figure 7.13: Calculated potential energy curves for increasing the edge length of the regular octagonal H_8 model with the cc-pVQZ basis.

Table 7.1: Comparison of equilibrium bond lengths and spectroscopic constants for some diatomic molecules. Basis set: cc-pV5Z, with correlation energy x^{-3}-extrapolated using cc-pVQZ and cc-pV5Z.

System	Method	R_e/Å	ω_e/cm^{-1}	$\omega_e x_e$/cm^{-1}
HF	CCSD	0.914	4198.7	93.9
	BCCD	0.914	4197.9	87.6
	BQVCCD	0.914	4200.9	87.9
	OQVCCD	0.914	4201.3	89.5
	CCSD(T)	0.917	4148.4	95.0
	BCCD(T)	0.917	4141.3	87.7
	BQVCCD(T)	0.917	4146.5	88.5
	OQVCCD(T)	0.917	4150.8	91.9
	Empirical[116]	0.917	4138.3	89.9
F_2	CCSD	1.388	1025.5	8.7
	BCCD	1.386	1030.3	8.7
	BQVCCD	1.385	1034.3	8.7
	OQVCCD	1.385	1034.6	8.6
	CCSD(T)	1.410	929.2	11.4
	BCCD(T)	1.410	933.3	11.3
	BQVCCD(T)	1.407	942.1	11.2
	OQVCCD(T)	1.407	942.9	11.2
	Empirical	1.412	916.6	11.2
N_2	CCSD	1.092	2445.3	12.8
	BCCD	1.091	2456.2	12.6
	BQVCCD	1.090	2464.0	12.5
	OQVCCD	1.090	2461.0	12.5
	CCSD(T)	1.099	2364.9	13.8
	BCCD(T)	1.099	2370.5	13.7
	BQVCCD(T)	1.097	2384.5	13.4
	OQVCCD(T)	1.098	2382.0	13.5
	Empirical	1.098	2358.6	14.3

7.4 The Surprising Implications of the Variational Upper Bound Property

The OQVCCD(T) method has been demonstrated to be extraordinarily robust for the prediction of potential energy curves involving the dissociation of multiply-bonded molecules, remaining in excellent quantitative agreement with MRCI+Q, and hence with experiment, across entire potential energy curves. Furthermore, it has been pointed out on several occasions that the onset of the catastrophic performance of CCSD(T) is coincident with the divergence of CCSD from OQVCCD, which is itself indicative of the divergence of CCSD from VCCSD. The contrast between the excellent performance of OQVCCD(T) and the problematic unphysical overestimation of the correlation energy by CCSD(T) in these examples provides compelling evidence that it is the non-variational character of CCSD itself that instigates the failure of CCSD(T), and *not* a breakdown of perturbation theory, to which it is commonly attributed. Put simply, the (T) correction does not cause the breakdown of CCSD(T), but instead simply magnifies the problematic lack of upper bound character inherent to CCSD in the multiply-bonded examples presented in the previous section.

These results therefore give new insight into the physics of the electron correlation problem, because it appears that a Coupled-Cluster-like method that possesses even an approximately-fulfilled upper bound property can adequately deal with even severe non-dynamic correlation across a range of systems, and when this occurs, the remaining effects of dynamic correlation may legitimately be included perturbatively. A pertinent question, of course, is why the (T) correction should be so sensitive to the presence of an upper bound property on the energy at all? It is, after all, a property that perturbation theory itself destroys; although VCCSD yields a rigorous upper bound on the exact ground-state Schrödinger energy eigenvalue, this would not be true of even a hypothetical VCCSD(T) theory.

The conclusion must be that it is not the upper bound property itself, but the more robust treatment of the electronic structure that this property enforces that

allows (T) to remain valid for OQVCCD when it fails for CCSD and BCCD. This makes sense because the only difference between the (T) correction to OQVCCD and the (T) correction to BCCD is in the values taken by the cluster amplitude parameters; the functional form of each is identical, as discussed above; the variational upper bound property on a calculated energy essentially constrains the values of the cluster amplitudes from which the energy is constructed. In fact, the difference between the cluster amplitudes calculated by OQVCCD and CCSD or BCCD is quite apparent in Figure 7.9, in which the square norm of the amplitudes is plotted as a function of the bond length of C_2. By constraining and stabilizing the values taken by the cluster amplitudes, the (T) correction itself inherits additional stability from the variational upper bound property.

It is, of course, still true that the (T) correction can break down when the highest occupied and lowest unoccupied molecular orbitals become degenerate in energy, and there is also the important point that more advanced perturbative corrections, such as (2)[80, 81, 111], exist that more effectively deal with triples (and quadruples) for non-dynamic cases by further decoupling the perturbative correction from the Hartree-Fock approximation. However, the variational upper bound property appears to greatly extend the range of systems for which single-reference *ab initio* electron correlation methods involving even the standard (T) perturbative correction can be appropriate. In particular, it is simply striking that the OQVCCD(T) method is capable of treating, and with both physical validity and quantitative accuracy, the complete potential energy curve of systems such as acetylene and dicarbon from the repulsive regime, through the equilibrium geometry and to the dissociation limit, and, to the author's knowledge, at the time of writing no other $\mathcal{O}(o^2v^4)$ iterative plus $\mathcal{O}(o^3v^4)$ non-iterative method (based on Restricted Hartree-Fock theory) has treated N_2 with accuracy even comparable to OQVCCD(T).

Chapter 8

Application to the Non-Linear Optical Properties of Model Hydrogen Chains[†]

In previous chapters, it has been discussed how the QVCCD ansatz inherits pseudo-variational upper bound character from its parent theory, VCCD, such that it is more robust to the breakdown of the Hartree-Fock approximation than methods based on Traditional Coupled Cluster theory. It was further demonstrated that, when combined with variational optimization of the orbitals to account for the effects of single excitations and a standard perturbative correction for triples, the resulting OQVCCD(T) method is capable of predicting both a physically correct and quantitatively accurate potential energy curve for the dissociation of dinitrogen, N_2, and other strongly-correlated molecules for which CCSD(T) and related methods fail catastrophically. In this chapter, a further pilot application of the Quasi-Variational Coupled Cluster method to the energies, polarizabilities and second hyperpolarizabilities of model hydrogen chains is presented, indicating it to be a suitable method also for the black-box treatment

[†]Relevant publication:

[119] J. B. Robinson and P. J. Knowles, J. Chem. Phys. (2012), *"Application of the Quasi-Variational Coupled Cluster Method to the Non-Linear Optical Properties of Model Hydrogen Systems"*, in press.

of multiradicals.

8.1 Motivation

The non-linear optical (NLO) properties of a molecule are responsible for macroscopic phenomena such as refractive indices[120], and the design of materials exhibiting large or systematically tunable NLO properties is of great interest for use in both current and future generations of optoelectronic devices[121–126]. Owing to this interest, a great deal of experimental[127–131] and theoretical[129–133] work has been targeted at this area. Linearly conjugated organic polymer chains have been extensively studied due to the discovery of relationships between structural and NLO properties[134–141], and, for singlet multiradicals in particular, studies on models have found the second hyperpolarizability to depend strongly on the diradical character[142, 143]. This has been confirmed theoretically for systems such as graphene nanoflakes[144, 145], as well as experimentally[146, 147].

The first investigation of molecular NLO properties using highly correlated wavefunction methods was performed by Bartlett and Purvis[148], and a correct treatment of electron correlation has since been shown to be essential in obtaining quantitatively accurate values of NLO properties from theoretical methods[149–157], especially second hyperpolarizabilities, which Hartree-Fock theory is known to systematically underestimate. Unfortunately, the common variants of Density Functional Theory[158, 159] (DFT) overestimate these same NLO properties[160, 161], and while more novel approaches have mitigated this problem, it has yet to be fully resolved[162, 163].

Due to the size and complexity of many of the systems of interest, *ab initio* calculations are typically applied to smaller representative test cases[143] or else resort to the investigation of model hydrogen chains as prototypical multiradicals[142, 164–169] in order to benchmark other, less expensive methods[150, 155, 170, 171]. However, since the NLO properties are extremely sensitive to the treatment of electron correlation[149, 172, 173], it has become important also to benchmark the *ab initio* methods amongst themselves on these systems[151, 156]. For this

purpose, it has become common to use Traditional Coupled Cluster limited to single and double excitations of the HF reference determinant in order to achieve a highly correlated approximation to the many-body Schrödinger wavefunction, sometimes combined with a correction to the calculated energy motivated by perturbative estimates of the effect of triple excitations.

Unfortunately, as has been discussed extensively throughout earlier chapters, TCC itself performs poorly when the Hartree-Fock approximation on which it is based becomes qualitatively wrong and a single-determinantal reference wavefunction becomes a poor model of the exact electronic wavefunction. This is especially problematic for multiradicals and geometries far from equilibrium, which are often the systems of interest in the context of applications to non-linear optics[143]. If the TCC energy becomes poor as a result of this problem, the effect will be magnified for higher-order properties, such as second hyperpolarizabilities. Furthermore, the multireference methods, while more reliable in such situations, do not operate in a black-box fashion, as discussed previously, and are thus more difficult to deploy on such large systems.

Therefore, it would be highly desirable if a robust and accurate single-reference method could be found for application to these systems, and, in this chapter, the performance of the Quasi-Variational Coupled Cluster method is tested for this task.

8.2 Methodology

Zero-frequency non-linear optical properties may be calculated from quantum-chemical methods by finite-difference differentiation of the calculated energies with respect to the strength of a small applied field; if the molecular Hamiltonian is perturbed by the application of a weak electric field, $\mathcal{F}$, the total energy, E, of the molecule may be written as a Taylor series in orders of the field strength[58].

$$E = \sum_{n=0}^{\infty} \frac{\mathcal{F}^n}{n!} \left. \frac{d^n E}{d\mathcal{F}^n} \right|_{\mathcal{F}=0} \tag{8.1}$$

The dipole moment of the molecule in the direction of the applied field, a measure of the separation of charge in the molecule along that axis, is,

$$\langle \mu \rangle = -\frac{dE}{d\mathcal{F}} = -\sum_{n=1}^{\infty} \frac{\mathcal{F}^{n-1}}{(n-1)!} \left. \frac{d^n E}{d\mathcal{F}^n} \right|_{\mathcal{F}=0} . \tag{8.2}$$

The first term in this series is the static (or permanent) dipole moment, and the other terms represent the contributions to the induced dipole moment. The coefficients of the second, third and fourth terms, which measure the response of the molecule to an applied field, are the electric polarizability, α, the hyperpolarizability, β, and the second hyperpolarizability, γ.

$$\alpha = -\left. \frac{d^2 E}{d\mathcal{F}^2} \right|_{\mathcal{F}=0} \tag{8.3}$$

$$\beta = -\frac{1}{2!} \left. \frac{d^3 E}{d\mathcal{F}^3} \right|_{\mathcal{F}=0} \tag{8.4}$$

$$\gamma = -\frac{1}{3!} \left. \frac{d^4 E}{d\mathcal{F}^4} \right|_{\mathcal{F}=0} . \tag{8.5}$$

In the following sections, results are presented that benchmark the performance of QVCC against CCSD, CCSD(T), BCCD and BCCD(T) on several model hydrogen systems, using either Full Configuration Interaction (FCI) or Multireference Averaged Quadratic Coupled Cluster[32] (MRAQCC) as the standard for correct behaviour in each case. The sensitivity of the NLO properties, in particular, to the treatment of electron correlation in these systems make them perfect tests of any new quantum-chemical method. In each case, a Restricted Hartree-Fock (RHF) reference wavefunction is used. Although better results could possibly be achieved by the use of an Unrestricted Hartree-Fock (UHF) reference wavefunction that correctly describes molecular dissociation, by instead using RHF in each case, a fair comparison between each of the methods can be achieved. The effectiveness of the RHF-based and UHF-based quantum-chemical methods is compared and contrasted separately in section 8.9. Additionally, in order to further establish a fair benchmark of comparison between the methods, the optimization of the orbitals in the QVCC calculations is substituted for the Brueckner condition, so that BQVCCD and BQVCCD(T) may be compared directly with BCCD and BCCD(T).

Each model system is assigned two degrees of freedom, and, in each case,

insight is gained into how the properties of interest vary through the examination of 1-D cuts of the potential energy surfaces or the surfaces of the polarizability or second hyperpolarizability. For several of the systems the accuracy of each of the single-reference electronic structure methods is then systematically gauged by obtaining a set of errors (computed by taking the value of the property of interest and subtracting from it the value obtained from a more accurate calculation, such as FCI) over a representative region of the corresponding potential energy surface. From each set of errors, $\{\epsilon_i\}$ with $i \in \{1, 2, \ldots, N\}$, three measures of the average error are computed; the mean signed error,

$$\bar{\epsilon} = \frac{1}{N} \sum_{i=1}^{N} \epsilon_i \tag{8.6}$$

the mean absolute error,

$$\bar{|\epsilon|} = \frac{1}{N} \sum_{i=1}^{N} |\epsilon_i| \tag{8.7}$$

and the root mean square error,

$$\sqrt{\bar{\epsilon^2}} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} {\epsilon_i}^2}, \tag{8.8}$$

and a measure of the spread of the errors, the standard deviation of the signed errors,

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\epsilon_i - \bar{\epsilon}\right)^2}. \tag{8.9}$$

In addition, a further indication of the distribution of errors is contained in the maximum (or most positive) error, $\lceil \epsilon \rceil$, and the minimum (or most negative) error, $\lfloor \epsilon \rfloor$, which, particularly for the case of the calculated energies, also describe the degree to which the upper bound property is in effect. The average error measures are taken to indicate the accuracy of the method, while the measures of the spread are taken to indicate the reliability (or consistency) of the method, such that a smaller spread indicates that whatever level of accuracy is achieved varies by a smaller amount over the potential energy surface. A supporting selection of FCI and MRAQCC energies, polarizabilities and second hyperpolarizabilities for each

Table 8.1: Calculated energies, polarizabilities and second hyperpolarizabilities for a selection of geometries of the various model hydrogen systems with the aug-cc-pVDZ basis, quoted in atomic units, and calculated with either the FCI (H_4) or MRAQCC (H_6) methods.

$D_{\infty h}$ H_4	R_1/Å	R_2/Å	E	α	γ
	1.00	1.00	-2.258545	$9.891\times10^{+0}$	$5.206\times10^{+1}$
	1.00	4.00	-2.283675	$1.158\times10^{+1}$	$5.107\times10^{+1}$
	2.00	1.75	-2.048722	$1.615\times10^{+1}$	$8.584\times10^{+1}$
	2.50	2.50	-2.008549	$1.673\times10^{+1}$	$1.072\times10^{+2}$
	3.00	4.00	-1.999858	$1.688\times10^{+1}$	$1.055\times10^{+2}$
C_{2v} H_4	R/Å	θ/°	E	α	γ
	2.25	0.0	-2.020093	$1.593\times10^{+1}$	$8.977\times10^{+1}$
	2.50	0.0	-2.009633	$1.626\times10^{+1}$	$1.016\times10^{+2}$
	2.75	0.0	-2.003707	$1.643\times10^{+1}$	$1.070\times10^{+2}$
	3.00	0.0	-2.000577	$1.653\times10^{+1}$	$1.075\times10^{+2}$
	3.25	0.0	-1.998985	$1.661\times10^{+1}$	$1.055\times10^{+2}$
D_{2h} H_4	R/Å	θ/°	E	α	γ
	1.75	80.0	-2.017706	$1.632\times10^{+1}$	$9.883\times10^{+1}$
	2.00	70.0	-2.016094	$1.668\times10^{+1}$	$1.050\times10^{+2}$
	2.00	80.0	-2.005927	$1.655\times10^{+1}$	$1.078\times10^{+2}$
	2.25	70.0	-2.005764	$1.675\times10^{+1}$	$1.084\times10^{+2}$
	2.25	80.0	-2.000833	$1.664\times10^{+1}$	$1.074\times10^{+2}$
C_{2h} H_4	R/Å	θ/°	E	α	γ
	1.0	45.0	-2.253203	$9.792\times10^{+0}$	$7.684\times10^{+1}$
	1.5	45.0	-2.130547	$1.441\times10^{+1}$	$6.347\times10^{+1}$
	2.0	45.0	-2.040944	$1.641\times10^{+1}$	$8.920\times10^{+1}$
	2.5	45.0	-2.008522	$1.671\times10^{+1}$	$1.068\times10^{+2}$
	3.0	45.0	-2.000026	$1.676\times10^{+1}$	$1.056\times10^{+2}$
$D_{\infty h}$ H_6	R_1/Å	R_2/Å	E	α	γ
	1.0	1.2	-3.400172	$1.471\times10^{+1}$	$6.743\times10^{+1}$
	1.2	1.8	-3.325041	$1.895\times10^{+1}$	$7.127\times10^{+1}$
	1.4	1.6	-3.234402	$2.065\times10^{+1}$	$7.499\times10^{+1}$
	1.6	1.6	-3.165001	$2.217\times10^{+1}$	$7.866\times10^{+1}$
	1.8	1.8	-3.105677	$2.370\times10^{+1}$	$8.935\times10^{+1}$

of the models is supplied in order to allow the scale of the errors presented to be established. This is given in Table 8.1.

Although much of the interest in model hydrogen systems is in the NLO properties of linear chains parallel to the longitudinal axis, here, this constraint is relaxed and several different model hydrogen systems are investigated. However, only the perpendicular NLO properties of each of the model systems are first treated. This is done for several reasons. First, in each of the models tested, there is a unique axis perpendicular to the plane of the system, whereas, for several of the models, no such unique axis exists in the plane. Thus, the perpendicular polarizabilities yield a single representative test for each system. Second, the perpendicular non-linear optical properties are less sensitive to the applied field strength and are thus more amenable to calculation by finite-difference differentiation. Third, the polarizabilities perpendicular to a bond axis are interesting in their own right, since they give a measure of the shape of the electronic structure, and, in particular, the contraction of the electron distribution as atoms are brought together to form covalent bonds. Finally, the goal of this chapter is only to present further evidence that the new and robust Quasi-Variational Coupled Cluster electronic structure ansatz allows the more accurate prediction of NLO properties when TCC fails to perform adequately, making it potentially valuable for future investigations. Perpendicular properties suffice for this.

Four of the model systems tested are illustrated in Figure 8.1, which are, in clockwise order from the top-left, $D_{\infty h}$, C_{2v}, D_{2h} and C_{2h} arrangements of four hydrogen atoms. Also investigated is the 6-atom equivalent of the $D_{\infty h}$ model. It is further established that the findings extend to parallel polarizabilities and hyperpolarizabilities by investigating the metal-insulator transition in $D_{\infty h}$ H_{10} as the nearest-neighbour separation of the atoms is increased.

8.3 The $D_{\infty h}$ H_4 Model

The $D_{\infty h}$ H_4 model consists of four hydrogen atoms arranged linearly, with R_1 corresponding to the distance between the outer and inner atoms, and with R_2

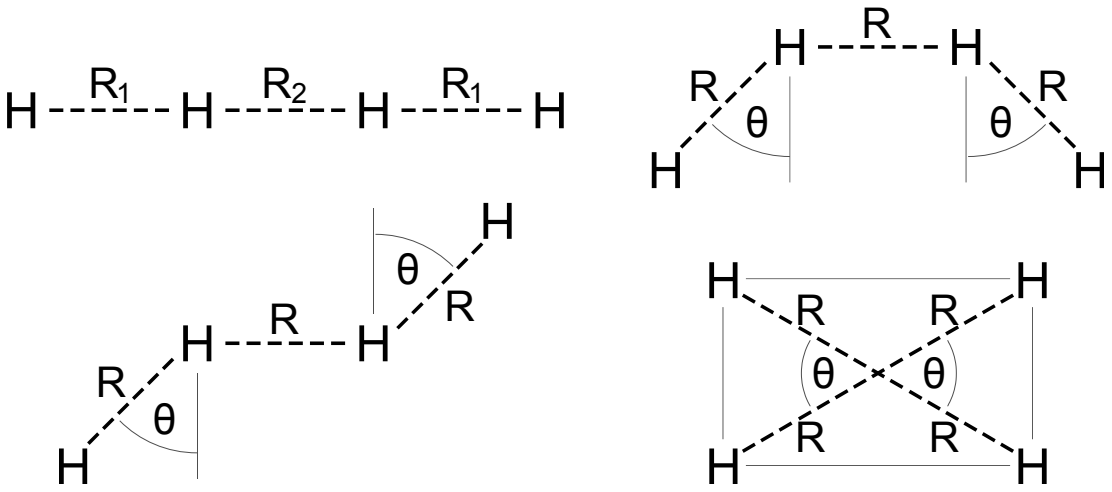


Figure 8.1: Clockwise from top-left, the $D_{\infty h}$, C_{2v}, D_{2h} and C_{2h} H_4 models.

corresponding to the distance between the two inner atoms, as is illustrated by the top-left diagram of Figure 8.1. For large R_2, the system corresponds to two isolated hydrogen molecules, each with bond length R_1, whereas for large R_1, the outer hydrogen atoms become isolated, leaving an inner hydrogen molecule with bond length R_2.

In Figure 8.2 it is shown how the energies vary for a uniform ($R_1 = R_2$) arrangement of the atoms as the common bond length is increased. In the case of the energy, BQVCCD clearly out-performs both CCSD and BCCD; the BQVCCD curve remains above FCI throughout, and the peak BQVCCD error, located around 2.2Å, is significantly smaller in magnitude than the TCC peak errors. The BQVCCD curve also remains fairly parallel to the FCI curve, and the error therefore remains uniform throughout, whereas both CCSD and BCCD display large fluctuations in accuracy relative to FCI. Furthermore, it is apparent that the addition of the triples correction to BQVCCD improves the overall description of the potential energy curve, resulting in smaller errors throughout, whereas it has the opposite effect on the TCC methods, magnifying their errors.

The error analysis for the $D_{\infty h}$ H_4 model is given in Table 8.2, and supports the observations already made. BQVCCD(T) possesses the smallest mean signed, mean absolute and root mean square errors of the methods tested, and by roughly an order of magnitude in some cases. It is noteworthy that the lowest (or most negative) BQVCCD energy error is, in fact, positive, with a value

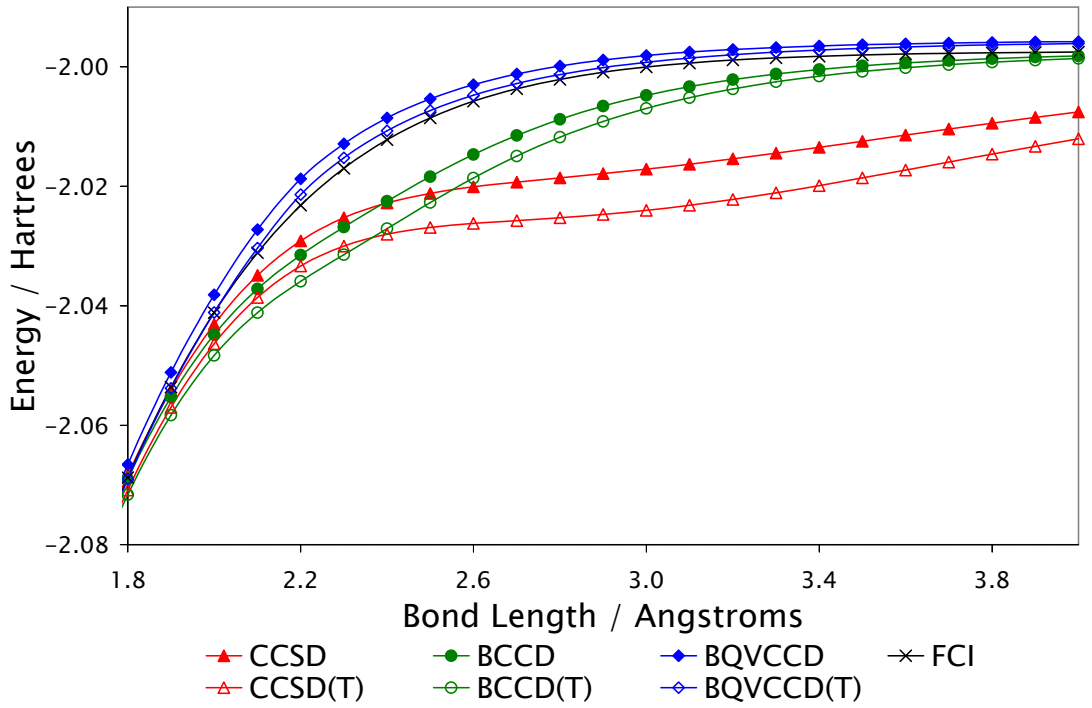


Figure 8.2: Calculated energies of the $D_{\infty h}$ H_4 model with $R_1 = R_2$ and the aug-cc-pVDZ basis.

of just 9 microHartrees, indicating that the approximately-fulfilled upper-bound property of the BQVCCD approach is in action here. The standard deviation of the BQVCCD(T) energy errors is also the smallest of any of the methods, indicating that BQVCCD(T) maintains roughly the same level of accuracy over the potential energy surface, whereas the TCC methods are far less predictable or reliable. The polarizability errors are similarly in favour of BQVCCD(T). The BQVCCD(T) mean signed error is, at 50 μa.u., at least two orders of magnitude smaller than any of the TCC values, and the mean absolute and root mean square errors are between 3 and 4 times smaller than the values predicted by BCCD, the best of the TCC methods in this case. The second hyperpolarizabilities are more interesting, however. For this property, the (T) correction does very little, and, in fact, makes the BQVCCD(T) root mean square error slightly worse, by approximately 0.14 a.u. However, of all the methods tested, BQVCCD still predicts the second hyperpolarizabilities closest in value to FCI, with a root mean square error roughly 6 times smaller than both CCSD and BCCD.

Table 8.2: Errors relative to FCI for calculated energies of, and polarizabilities perpendicular to the $D_{\infty h}$ H_4 model with the aug-cc-pVDZ basis. Results were obtained from the set of points $\{(R_1, R_2)\}$ where $R_1 \in \{1.0, 1.5, 2.0, 2.5, 3.0\}$Å and $R_2 \in \{1.0, 1.75, 2.5, 3.25, 4.0\}$Å.

		CCSD	CCSD(T)	BCCD	BCCD(T)	BQVCCD	BQVCCD(T)
E	$\bar{\epsilon}$	-3.42×10^{-3}	-6.61×10^{-3}	-1.20×10^{-3}	-2.96×10^{-3}	1.87×10^{-3}	5.24×10^{-4}
	$\lvert\bar{\epsilon}\rvert$	4.49×10^{-3}	6.65×10^{-3}	2.33×10^{-3}	3.01×10^{-3}	1.87×10^{-3}	7.07×10^{-4}
	$\sqrt{\bar{\epsilon^2}}$	9.97×10^{-3}	1.47×10^{-2}	5.48×10^{-3}	7.87×10^{-3}	2.94×10^{-3}	1.21×10^{-3}
	$\lceil\epsilon\rceil$	3.60×10^{-3}	2.31×10^{-4}	3.50×10^{-3}	2.31×10^{-4}	9.49×10^{-3}	3.98×10^{-3}
	$\lfloor\epsilon\rfloor$	-4.16×10^{-2}	-5.80×10^{-2}	-2.33×10^{-2}	-3.31×10^{-2}	9.36×10^{-6}	-2.05×10^{-3}
	σ	9.56×10^{-3}	1.34×10^{-2}	5.45×10^{-3}	7.44×10^{-3}	2.31×10^{-3}	1.11×10^{-3}
α	$\bar{\epsilon}$	-1.80×10^{-1}	-2.36×10^{-1}	1.43×10^{-2}	3.25×10^{-2}	-3.04×10^{-3}	-5.76×10^{-5}
	$\lvert\bar{\epsilon}\rvert$	1.82×10^{-1}	2.49×10^{-1}	1.02×10^{-1}	1.24×10^{-1}	4.17×10^{-2}	3.06×10^{-2}
	$\sqrt{\bar{\epsilon^2}}$	4.29×10^{-1}	5.89×10^{-1}	2.95×10^{-1}	3.74×10^{-1}	8.71×10^{-2}	6.97×10^{-2}
	$\lceil\epsilon\rceil$	6.51×10^{-3}	9.18×10^{-2}	$1.19\times10^{+0}$	$1.54\times10^{+0}$	3.46×10^{-1}	2.88×10^{-1}
	$\lfloor\epsilon\rfloor$	$-1.50\times10^{+0}$	$-2.18\times10^{+0}$	-8.46×10^{-1}	$-1.03\times10^{+0}$	-2.06×10^{-1}	-1.68×10^{-1}
	σ	3.97×10^{-1}	5.51×10^{-1}	3.00×10^{-1}	3.81×10^{-1}	8.89×10^{-2}	7.12×10^{-2}
γ	$\bar{\epsilon}$	6.56×10^{-1}	-7.40×10^{-1}	$2.06\times10^{+0}$	$1.25\times10^{+0}$	7.43×10^{-1}	4.53×10^{-1}
	$\lvert\bar{\epsilon}\rvert$	$4.43\times10^{+0}$	$5.24\times10^{+0}$	$3.74\times10^{+0}$	$3.67\times10^{+0}$	$1.12\times10^{+0}$	$1.11\times10^{+0}$
	$\sqrt{\bar{\epsilon^2}}$	$9.64\times10^{+0}$	$1.09\times10^{+1}$	$1.02\times10^{+1}$	$1.00\times10^{+1}$	$1.63\times10^{+0}$	$1.77\times10^{+0}$
	$\lceil\epsilon\rceil$	$3.56\times10^{+1}$	$3.10\times10^{+1}$	$3.97\times10^{+1}$	$3.43\times10^{+1}$	$3.91\times10^{+0}$	$3.79\times10^{+0}$
	$\lfloor\epsilon\rfloor$	$-2.00\times10^{+1}$	$-2.76\times10^{+1}$	$-1.77\times10^{+1}$	$-2.55\times10^{+1}$	$-2.42\times10^{+0}$	$-5.21\times10^{+0}$
	σ	$9.82\times10^{+0}$	$1.11\times10^{+1}$	$1.02\times10^{+1}$	$1.02\times10^{+1}$	$1.48\times10^{+0}$	$1.74\times10^{+0}$

8.4 The C_{2v} H_4 Model

The C_{2v} model, illustrated by the diagram in the top-right of Figure 8.1, corresponds to a 'cis' arrangement of the four hydrogen atoms, with the degrees of freedom chosen to be the distance between nearest neighbours and the angle subtended by lines from an outer atom to an inner atom and along the perpendicular to the line joining the two inner atoms.

The most interesting degree of freedom to vary in this system is the angle θ; the outer hydrogen atoms are nearby for $\theta < 0°$, but as the system is distorted through $\theta = 0°$, corresponding to the square geometry, the outer atoms swing apart and the optimum Hartree-Fock reference wavefunction changes. This results in cusps in the potential energy curves predicted by the single-reference methods at $\theta = 0°$, as can be seen in Figure 8.3. The CCSD and BCCD methods level off too quickly, forming concave cusps, whereas the BQVCCD cusp is convex and the curve mimics the FCI shape more correctly overall. It is again true that the (T) correction to the BQVCCD method results in an improved curve, even

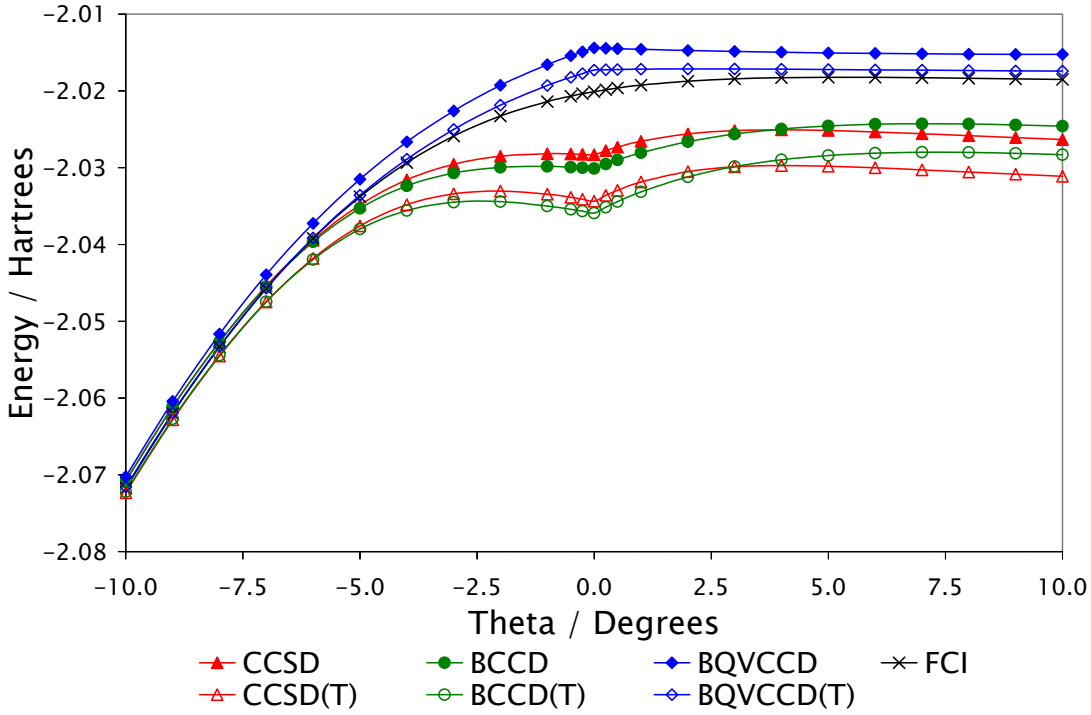


Figure 8.3: Calculated potential energy curves for the C_{2v} H_4 model with $R = 2.25$Å, and with the aug-cc-pVDZ basis.

flattening the cusp, whereas the (T) correction to the TCC methods pushes the curves further from FCI quality and sharpen the cusps.

The error analysis for C_{2v} H_4 is given in Table 8.3. As can be expected from the plot, the BQVCCD(T) method possesses the smallest errors in the calculated energies across all the main categories. It is particularly striking that the BQVCCD(T) root mean square error, at 0.7 milliHartrees, is 20 times smaller than the CCSD(T) error, at 14 milliHartrees, and 11 times smaller than the BCCD(T) value, at 8 milliHartrees.

The polarizability errors are also quite impressive; the BQVCCD(T) mean absolute error for the perpendicular polarizability is 0.025 a.u., whereas the CCSD(T) and BCCD(T) errors are larger by factors of approximately 13 and 5 respectively. The standard deviations of the polarizability errors are also roughly an order of magnitude smaller for BQVCCD and BQVCCD(T) than for the TCC methods. The BCCD method predicts just slightly better second hyperpolarizability values than BQVCCD for this system, as measured by mean absolute and root mean square errors. However, the standard deviation of the errors indi-

Table 8.3: Errors relative to FCI for calculated energies, and polarizabilities perpendicular to the plane, in the C_{2v} H_4 model with the aug-cc-pVDZ basis. Results were obtained from the set of points $\{(R, \theta)\}$ where $R \in \{2.25, 2.5, 2.75, 3.0, 3.25\}$Å and $\theta \in \{0, \pm 2, \pm 4, \pm 6, \pm 8\}^\circ$.

		CCSD	CCSD(T)	BCCD	BCCD(T)	BQVCCD	BQVCCD(T)
E	$\bar{\epsilon}$	-7.14×10^{-3}	-1.12×10^{-2}	-4.28×10^{-3}	-6.81×10^{-3}	1.65×10^{-3}	1.68×10^{-4}
	$\lvert\bar{\epsilon}\rvert$	7.17×10^{-3}	1.12×10^{-2}	4.32×10^{-3}	6.81×10^{-3}	1.65×10^{-3}	4.35×10^{-4}
	$\sqrt{\bar{\epsilon^2}}$	9.34×10^{-3}	1.38×10^{-2}	5.04×10^{-3}	7.73×10^{-3}	1.99×10^{-3}	6.75×10^{-4}
	$\lceil\epsilon\rceil$	6.52×10^{-4}	-1.13×10^{-3}	6.64×10^{-4}	-1.08×10^{-3}	5.69×10^{-3}	2.80×10^{-3}
	$\lfloor\epsilon\rfloor$	-2.17×10^{-2}	-3.07×10^{-2}	-1.00×10^{-2}	-1.58×10^{-2}	5.77×10^{-4}	-5.51×10^{-4}
	σ	6.10×10^{-3}	8.18×10^{-3}	2.68×10^{-3}	3.70×10^{-3}	1.12×10^{-3}	6.62×10^{-4}
α	$\bar{\epsilon}$	-1.45×10^{-1}	-2.48×10^{-1}	1.87×10^{-2}	4.33×10^{-2}	-3.25×10^{-2}	-1.89×10^{-2}
	$\lvert\bar{\epsilon}\rvert$	2.32×10^{-1}	3.29×10^{-1}	9.78×10^{-2}	1.18×10^{-1}	3.26×10^{-2}	2.48×10^{-2}
	$\sqrt{\bar{\epsilon^2}}$	3.09×10^{-1}	4.57×10^{-1}	1.16×10^{-1}	1.41×10^{-1}	4.35×10^{-2}	3.52×10^{-2}
	$\lceil\epsilon\rceil$	2.37×10^{-1}	2.50×10^{-1}	1.84×10^{-1}	2.37×10^{-1}	1.22×10^{-3}	3.70×10^{-2}
	$\lfloor\epsilon\rfloor$	-7.17×10^{-1}	$-1.04\times10^{+0}$	-2.23×10^{-1}	-2.66×10^{-1}	-1.26×10^{-1}	-1.07×10^{-1}
	σ	2.77×10^{-1}	3.88×10^{-1}	1.16×10^{-1}	1.35×10^{-1}	2.91×10^{-2}	3.00×10^{-2}
γ	$\bar{\epsilon}$	$3.32\times10^{+0}$	$5.11\times10^{+0}$	$-1.05\times10^{+0}$	$-2.28\times10^{+0}$	$1.75\times10^{+0}$	$1.72\times10^{+0}$
	$\lvert\bar{\epsilon}\rvert$	$5.23\times10^{+0}$	$6.84\times10^{+0}$	$1.57\times10^{+0}$	$2.60\times10^{+0}$	$1.77\times10^{+0}$	$1.79\times10^{+0}$
	$\sqrt{\bar{\epsilon^2}}$	$7.12\times10^{+0}$	$9.81\times10^{+0}$	$2.08\times10^{+0}$	$3.42\times10^{+0}$	$2.21\times10^{+0}$	$2.65\times10^{+0}$
	$\lceil\epsilon\rceil$	$1.46\times10^{+1}$	$1.98\times10^{+1}$	$1.85\times10^{+0}$	$1.56\times10^{+0}$	$6.06\times10^{+0}$	$9.35\times10^{+0}$
	$\lfloor\epsilon\rfloor$	$-4.40\times10^{+0}$	$-5.30\times10^{+0}$	$-5.43\times10^{+0}$	$-8.38\times10^{+0}$	-4.72×10^{-1}	-8.88×10^{-1}
	σ	$6.36\times10^{+0}$	$8.47\times10^{+0}$	$1.81\times10^{+0}$	$2.58\times10^{+0}$	$1.37\times10^{+0}$	$2.04\times10^{+0}$

cates that the BQVCCD method still has the smallest spread of errors, so that it remains the most reliable method for calculating second hyperpolarizabilities, despite its slightly poorer mean accuracy here.

8.5 The D_{2h} H_4 Model

Next, the D_{2h} H_4 model[44], shown in the bottom-right of Figure 8.1 is examined, in which four hydrogen atoms are arranged in a rectangle that can be defined by the parameters R, which controls the distance of each H atom from the centre of mass, and θ, the angle subtended at the centre of mass by radii to two neighbouring vertices of the rectangle. The system is symmetric about $\theta = 90^\circ$, and the optimum Hartree-Fock reference wavefunction differs on either side of this line. Thus, the Hartree-Fock approximation breaks down around $\theta = 90^\circ$ as two determinants become equally important to the description of the ground-state electronic structure, and this makes the system an excellent test of

single-reference post-Hartree-Fock electron correlation methods.

It has been previously shown that the QVCCD method mimics the behaviour of VCCD (and thus FCI) well in this system[95], and the improved shape of the potential energy curve relative to the TCC methods around the square geometry can be seen in Figure 8.4. However, the BQVCCD method additionally models the non-linear optical properties of this system extremely well, achieving the correct shape for the second hyperpolarizability around $\theta = 90°$ despite a displacement in the curve position, as can be seen in Figure 8.5. The cusp present in the BQVCCD and BQVCCD(T) potential energy curves has also almost vanished in the second hyperpolarizability curves. In contrast, the poor quality of the CCSD and BCCD methods in the interval $[80, 100]°$ that is apparent in the calculated energies deteriorates even further for this more challenging property, with CCSD, CCSD(T), BCCD and BCCD(T) all predicting curves with incorrect slope relative to FCI, and possessing even sharper cusps.

As measured by the mean absolute and root mean square error data presented in Table 8.4, it is clear that BQVCCD(T) predicts the potential energy curves and second hyperpolarizabilities that best approximate the FCI values overall, and that BQVCCD predicts the best polarizabilities. Similarly, the BQVCCD and BQVCCD(T) standard deviations are smaller than the equivalent TCC values, from which it is possible to infer that not only are the accuracies of BQVCCD-calculated energies, polarizabilities and second hyperpolarizabilities greatly improved, but that the calculations are also stabilized, resulting in more systematically predictable, consistent errors.

8.6 The C_{2h} H_4 Model

The C_{2h} model is a simple modification of the C_{2v} model such that one of the outer hydrogens is on the opposite side, forming a 'trans' structure. The diagram is given in the bottom-left of Figure 8.1. Increasing the angle θ causes the outer hydrogen atoms to swap to opposite sides, with $\theta = 90°$ corresponding to the linear geometry.

Table 8.4: Errors relative to FCI for calculated energies of, and polarizabilities perpendicular to the D_{2h} H_4 model with the aug-cc-pVDZ basis. Results were obtained from the set of points $\{(R, \theta)\}$ where $R \in \{1.0, 1.75, 2.0, 2.25\}$Å and $\theta \in \{70, 72, 74, 76, 78, 80, 82, 84, 86, 87, 88, 89\}°$.

		CCSD	CCSD(T)	BCCD	BCCD(T)	BQVCCD	BQVCCD(T)
E	$\bar{\epsilon}$	-7.69×10^{-4}	-2.59×10^{-3}	-1.09×10^{-3}	-2.83×10^{-3}	1.36×10^{-3}	2.89×10^{-5}
	$\lvert\bar{\epsilon}\rvert$	1.85×10^{-3}	2.63×10^{-3}	2.16×10^{-3}	2.89×10^{-3}	1.36×10^{-3}	3.01×10^{-4}
	$\sqrt{\bar{\epsilon}^2}$	2.60×10^{-3}	3.96×10^{-3}	3.04×10^{-3}	4.39×10^{-3}	1.79×10^{-3}	3.96×10^{-4}
	$\lceil\epsilon\rceil$	3.88×10^{-3}	1.78×10^{-4}	3.33×10^{-3}	2.28×10^{-4}	5.98×10^{-3}	9.36×10^{-4}
	$\lfloor\epsilon\rfloor$	-7.15×10^{-3}	-1.15×10^{-2}	-8.56×10^{-3}	-1.27×10^{-2}	5.98×10^{-4}	-1.50×10^{-3}
	σ	2.51×10^{-3}	3.03×10^{-3}	2.87×10^{-3}	3.39×10^{-3}	1.18×10^{-3}	3.99×10^{-4}
α	$\bar{\epsilon}$	4.36×10^{-2}	2.92×10^{-2}	8.18×10^{-3}	1.79×10^{-2}	-1.53×10^{-2}	-1.37×10^{-2}
	$\lvert\bar{\epsilon}\rvert$	6.33×10^{-2}	5.74×10^{-2}	4.71×10^{-2}	5.68×10^{-2}	2.35×10^{-2}	2.46×10^{-2}
	$\sqrt{\bar{\epsilon}^2}$	8.91×10^{-2}	8.95×10^{-2}	6.85×10^{-2}	8.68×10^{-2}	3.42×10^{-2}	3.74×10^{-2}
	$\lceil\epsilon\rceil$	2.24×10^{-1}	2.32×10^{-1}	1.71×10^{-1}	2.21×10^{-1}	2.43×10^{-2}	2.61×10^{-2}
	$\lfloor\epsilon\rfloor$	-1.58×10^{-1}	-2.03×10^{-1}	-1.77×10^{-1}	-2.16×10^{-1}	-1.05×10^{-1}	-1.30×10^{-1}
	σ	7.85×10^{-2}	8.55×10^{-2}	6.88×10^{-2}	8.58×10^{-2}	3.09×10^{-2}	3.51×10^{-2}
γ	$\bar{\epsilon}$	$-1.07\times10^{+0}$	-6.49×10^{-1}	-2.04×10^{-1}	-7.80×10^{-1}	9.81×10^{-1}	7.74×10^{-1}
	$\lvert\bar{\epsilon}\rvert$	$1.59\times10^{+0}$	$1.09\times10^{+0}$	9.08×10^{-1}	$1.27\times10^{+0}$	9.88×10^{-1}	8.18×10^{-1}
	$\sqrt{\bar{\epsilon}^2}$	$1.92\times10^{+0}$	$1.61\times10^{+0}$	$1.41\times10^{+0}$	$1.94\times10^{+0}$	$1.26\times10^{+0}$	$1.07\times10^{+0}$
	$\lceil\epsilon\rceil$	$3.94\times10^{+0}$	$2.32\times10^{+0}$	$3.68\times10^{+0}$	$2.27\times10^{+0}$	$2.48\times10^{+0}$	$2.16\times10^{+0}$
	$\lfloor\epsilon\rfloor$	$-4.01\times10^{+0}$	$-4.71\times10^{+0}$	$-3.75\times10^{+0}$	$-5.98\times10^{+0}$	-1.06×10^{-1}	-5.54×10^{-1}
	σ	$1.61\times10^{+0}$	$1.49\times10^{+0}$	$1.41\times10^{+0}$	$1.79\times10^{+0}$	7.92×10^{-1}	7.42×10^{-1}

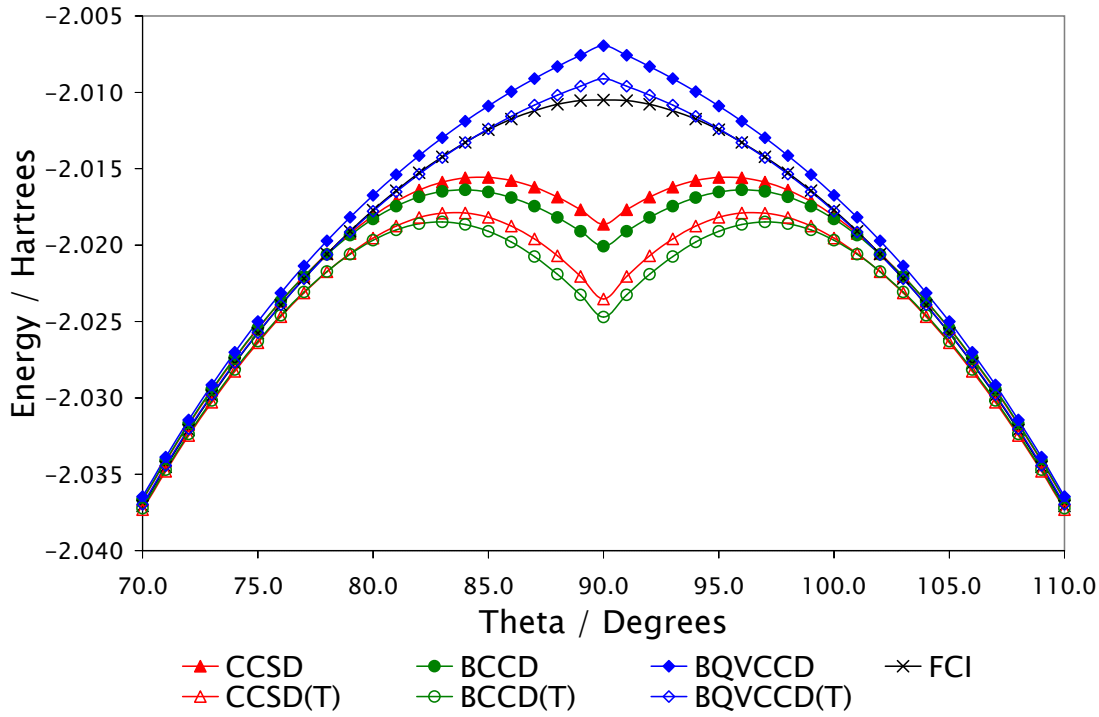


Figure 8.4: Calculated potential energy curves for the D_{2h} H_4 model with $R = 1.75$Å, and with the aug-cc-pVDZ basis.

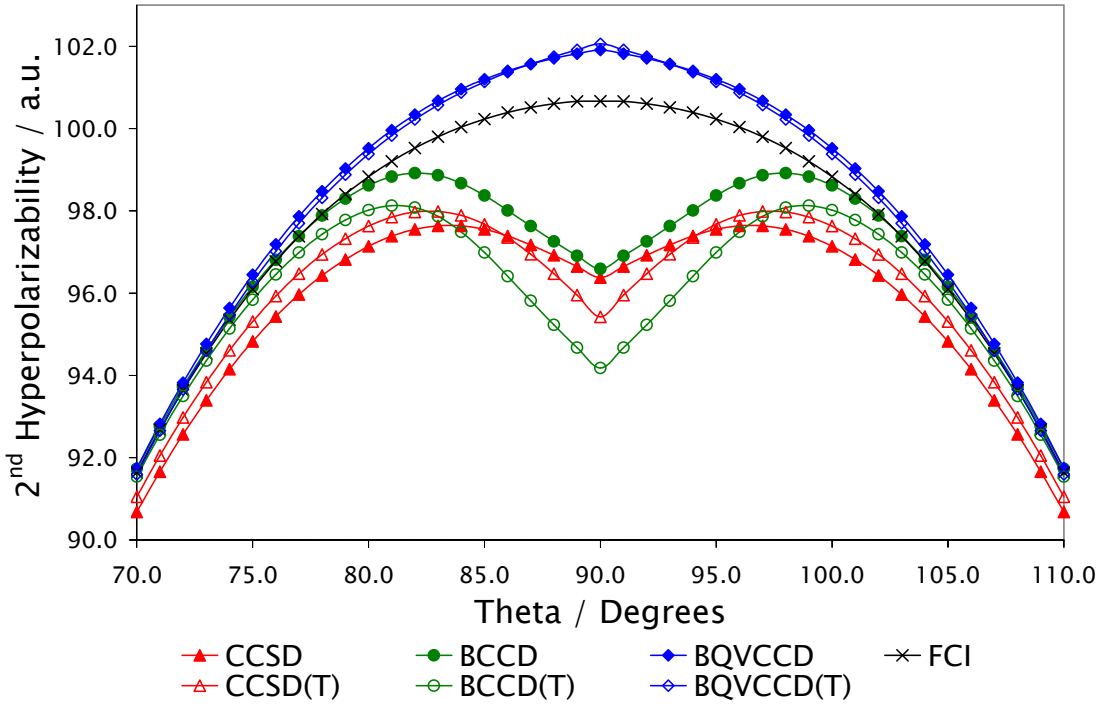


Figure 8.5: Calculated second hyperpolarizabilities perpendicular to the plane of the D_{2h} H_4 model with $R = 1.75$Å, and with the aug-cc-pVDZ basis

A plot of the potential energy curves obtained by varying the angle θ is shown in Figure 8.6. Both the CCSD and BCCD curves lie below FCI throughout, and the effect of the (T) correction is to push the curves lower still, further from FCI. However, the BQVCCD curve lies significantly above the FCI curve throughout, and the effect of the correction remains to push the energy down, resulting in the BQVCCD(T) curve being almost coincident with FCI. The calculated polarizabilities are also shown in Figure 8.7, and although each of the methods deteriorates in quality for the calculation of this more difficult property, BQVCCD(T) remains in extremely good agreement with FCI throughout.

These graphs add credence to the error analysis presented in Table 8.5, which indicate that BQVCCD(T) performs best for each of the properties and for almost all error measures. For example, the BQVCCD(T) root mean square energy error is 18 and 9 times smaller than the CCSD(T) and BCCD(T) errors respectively. The polarizabilities and second hyperpolarizabilities are similarly impressive, with errors 14 and 7 times smaller for the polarizabilities and by 5 and 4 times for the second hyperpolarizabilities respectively.

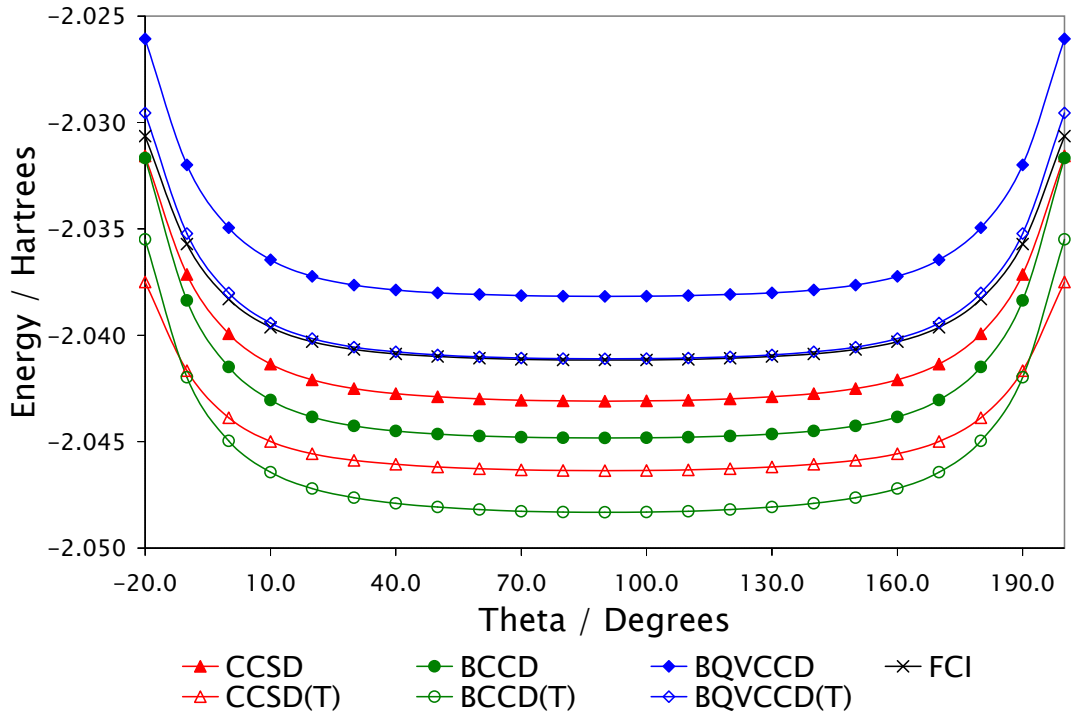


Figure 8.6: Calculated potential energy curves for the C_{2h} H_4 model with $R = 2.0$Å, and with the aug-cc-pVDZ basis.

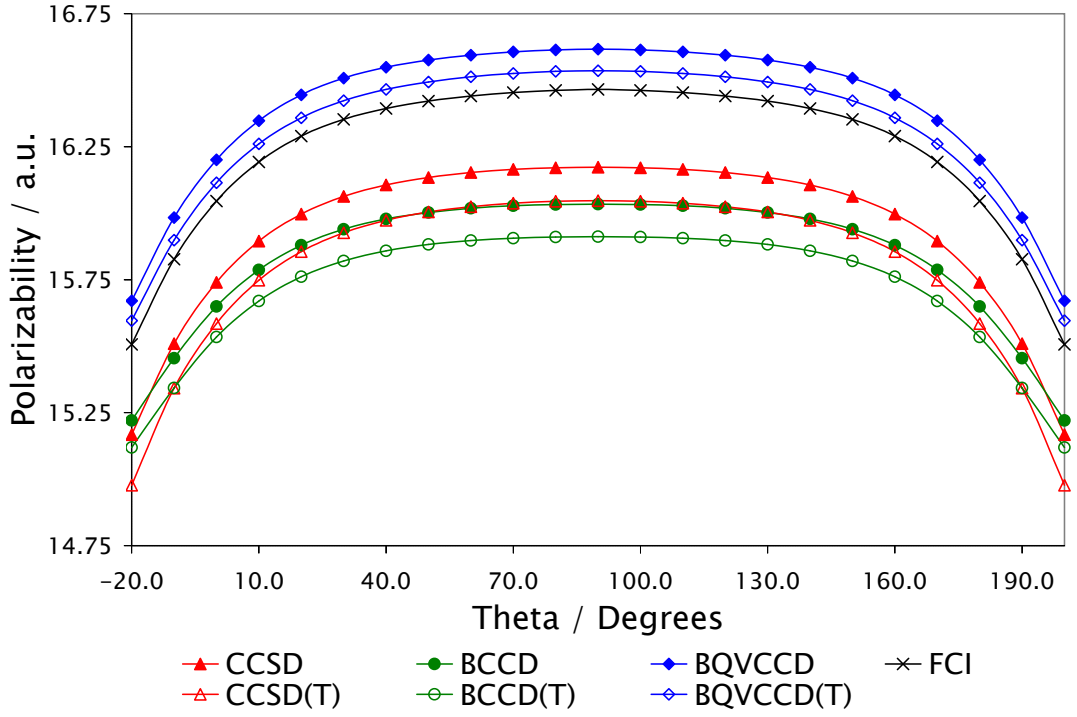


Figure 8.7: Calculated polarizabilities perpendicular to the plane of the C_{2h} H_4 model with $R = 2.0$Å, and with the aug-cc-pVDZ basis.

8.7 The $D_{\infty h}$ H_6 Model

Analogous to the $D_{\infty h}$ H_4 model, the more severe test case of the $D_{\infty h}$ H_6 model has additionally been studied, in which six hydrogen atoms are arranged linearly,

Table 8.5: Errors relative to FCI for calculated energies of, and polarizabilities perpendicular to the C_{2h} H_4 model with the aug-cc-pVDZ basis. Results were obtained from the set of points $\{(R, \theta)\}$ where $R \in \{1.0, 1.5, 2.0, 2.5, 3.0\}$Å and $\theta \in \{-15, 0, 15, 30, 45, 60, 75, 90\}^\circ$.

		CCSD	CCSD(T)	BCCD	BCCD(T)	BQVCCD	BQVCCD(T)
E	$\bar{\epsilon}$	-6.31×10^{-3}	-1.04×10^{-2}	-2.97×10^{-3}	-5.53×10^{-3}	2.45×10^{-3}	6.55×10^{-4}
	$\bar{\lvert\epsilon\rvert}$	7.54×10^{-3}	1.05×10^{-2}	4.13×10^{-3}	5.64×10^{-3}	2.45×10^{-3}	6.55×10^{-4}
	$\sqrt{\bar{\epsilon^2}}$	1.07×10^{-2}	1.53×10^{-2}	5.14×10^{-3}	7.62×10^{-3}	2.59×10^{-3}	8.55×10^{-4}
	$\lceil\epsilon\rceil$	2.59×10^{-3}	3.11×10^{-4}	2.24×10^{-3}	3.13×10^{-4}	4.05×10^{-3}	1.81×10^{-3}
	$\lfloor\epsilon\rfloor$	-2.38×10^{-2}	-3.39×10^{-2}	-9.90×10^{-3}	-1.42×10^{-2}	1.27×10^{-3}	5.71×10^{-5}
	σ	8.70×10^{-3}	1.13×10^{-2}	4.25×10^{-3}	5.30×10^{-3}	8.43×10^{-4}	5.57×10^{-4}
α	$\bar{\epsilon}$	-2.57×10^{-1}	-3.78×10^{-1}	-4.57×10^{-2}	-5.20×10^{-2}	1.31×10^{-2}	1.33×10^{-3}
	$\bar{\lvert\epsilon\rvert}$	2.58×10^{-1}	3.78×10^{-1}	1.33×10^{-1}	1.79×10^{-1}	5.23×10^{-2}	2.74×10^{-2}
	$\sqrt{\bar{\epsilon^2}}$	3.67×10^{-1}	5.24×10^{-1}	2.05×10^{-1}	2.70×10^{-1}	7.48×10^{-2}	3.75×10^{-2}
	$\lceil\epsilon\rceil$	1.59×10^{-2}	-3.14×10^{-3}	2.14×10^{-1}	3.00×10^{-1}	1.59×10^{-1}	7.67×10^{-2}
	$\lfloor\epsilon\rfloor$	-8.88×10^{-1}	$-1.24\times10^{+0}$	-4.31×10^{-1}	-5.53×10^{-1}	-6.18×10^{-2}	-5.21×10^{-2}
	σ	2.66×10^{-1}	3.68×10^{-1}	2.02×10^{-1}	2.68×10^{-1}	7.46×10^{-2}	3.80×10^{-2}
γ	$\bar{\epsilon}$	4.67×10^{-1}	-3.17×10^{-1}	7.54×10^{-1}	-2.97×10^{-1}	-2.03×10^{-1}	-2.84×10^{-1}
	$\bar{\lvert\epsilon\rvert}$	$6.87\times10^{+0}$	$9.08\times10^{+0}$	$6.14\times10^{+0}$	$7.63\times10^{+0}$	$2.14\times10^{+0}$	$1.95\times10^{+0}$
	$\sqrt{\bar{\epsilon^2}}$	$9.93\times10^{+0}$	$1.35\times10^{+1}$	$8.69\times10^{+0}$	$1.12\times10^{+1}$	$2.88\times10^{+0}$	$2.76\times10^{+0}$
	$\lceil\epsilon\rceil$	$1.48\times10^{+1}$	$1.73\times10^{+1}$	$1.44\times10^{+1}$	$1.57\times10^{+1}$	$2.85\times10^{+0}$	$2.40\times10^{+0}$
	$\lfloor\epsilon\rfloor$	$-2.16\times10^{+1}$	$-3.12\times10^{+1}$	$-1.77\times10^{+1}$	$-2.55\times10^{+1}$	$-6.11\times10^{+0}$	$-6.65\times10^{+0}$
	σ	$1.00\times10^{+1}$	$1.36\times10^{+1}$	$8.77\times10^{+0}$	$1.13\times10^{+1}$	$2.91\times10^{+0}$	$2.78\times10^{+0}$

with bond lengths alternating as R_1, R_2, R_1, R_2, R_1.

First, the line along the potential energy surface defined by $R_1 = R_2$ is investigated, for which the potential energy curves are given in Figure 8.8. All methods perform similarly well until approximately 1.6Å. At this bond length, the TCC methods begin to degrade significantly in quality, dropping below the MRAQCC energy in a non-variational fashion. However, the BQVCCD and BQVCCD(T) energies, supported by the approximately-fulfilled upper bound property, remain in excellent agreement with MRAQCC throughout and do not appear to degrade at all.

These findings extend also to the perpendicular polarizabilities, given in Figure 8.9, and second hyperpolarizabilities, given in Figure 8.10. The TCC methods predict a decreasing polarizability from approximately 1.8Å, which is clearly at odds with the smooth and monotonically increasing MRAQCC polarizability. The BQVCCD predicted polarizabilities, on the other hand, are far superior, and although the polarizability decreases around 2.2Å, this is quickly corrected such

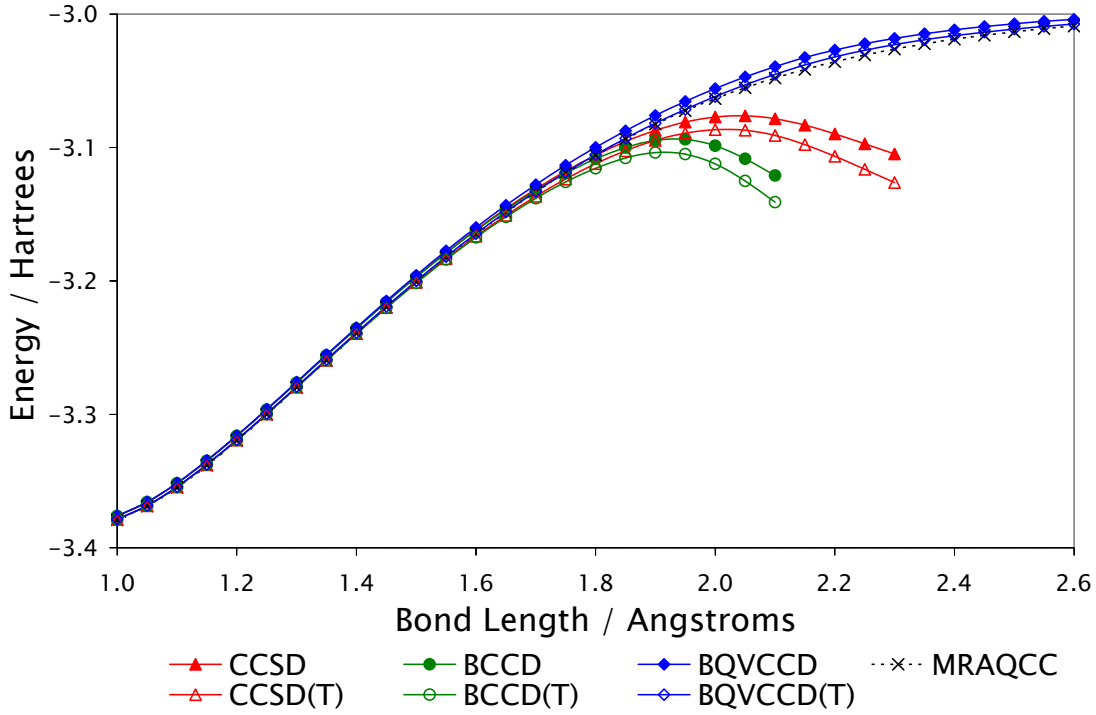


Figure 8.8: Calculated potential energy curves for the $D_{\infty h}$ H_6 model with $R_1 = R_2$ and with the aug-cc-pVDZ basis.

that the BQVCCD curve remains near the MRAQCC curve thoughout. The second hyperpolarizabilities are even more striking, with the TCC methods predicting values several times too large for 1.8-2.2Å. The BQVCCD curves again experience crossings with the MRAQCC curve, but the predicted values remain quantitatively accurate at all points.

The error analysis for this system is given in Table 8.6. An examination of the mean absolute errors for each of the three properties confirms that BQVCCD(T) is the most accurate of the single-reference methods, and by an order of magnitude in each case. The standard deviations also attest the reliability of the BQVCCD(T) method, with CCSD values factors of 5, 8 and 8 worse for the energy, polarizability and second hyperpolarizability respectively. Finally, once again, for this set of data, the addition of the (T) correction to the BQVCCD method yields smaller errors overall, whereas its addition to the TCC methods tends to have the opposite effect.

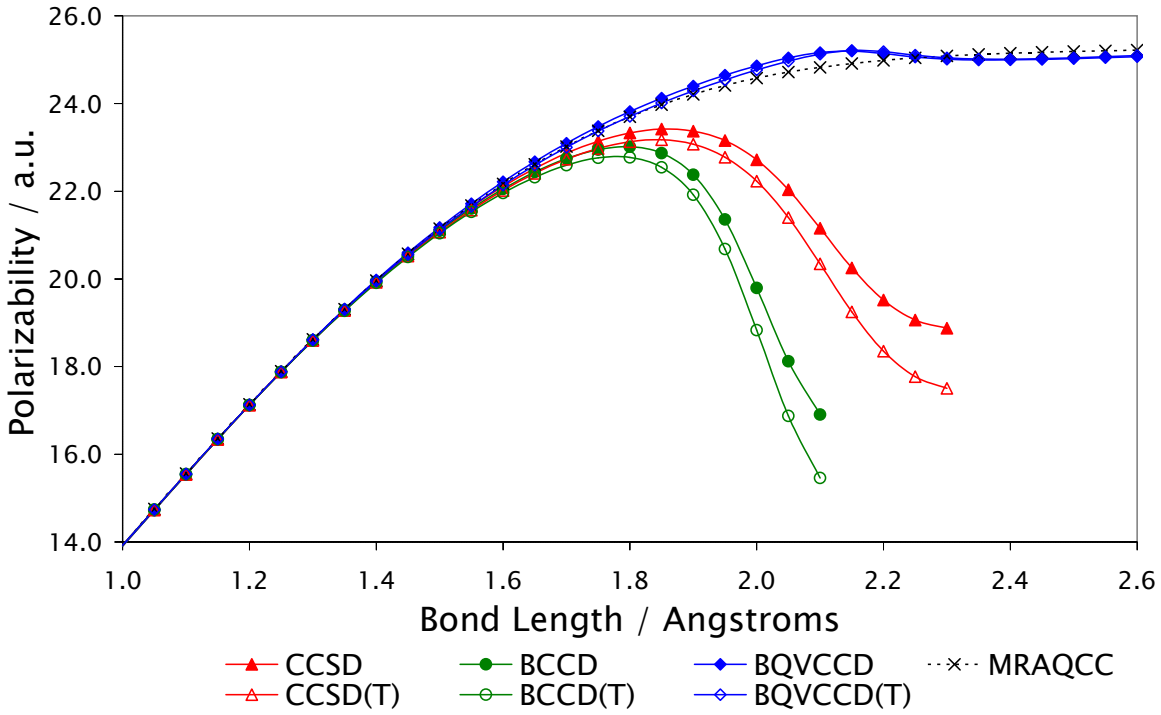


Figure 8.9: Calculated polarizabilities for the $D_{\infty h}$ H_6 model with $R_1 = R_2$ and with the aug-cc-pVDZ basis.

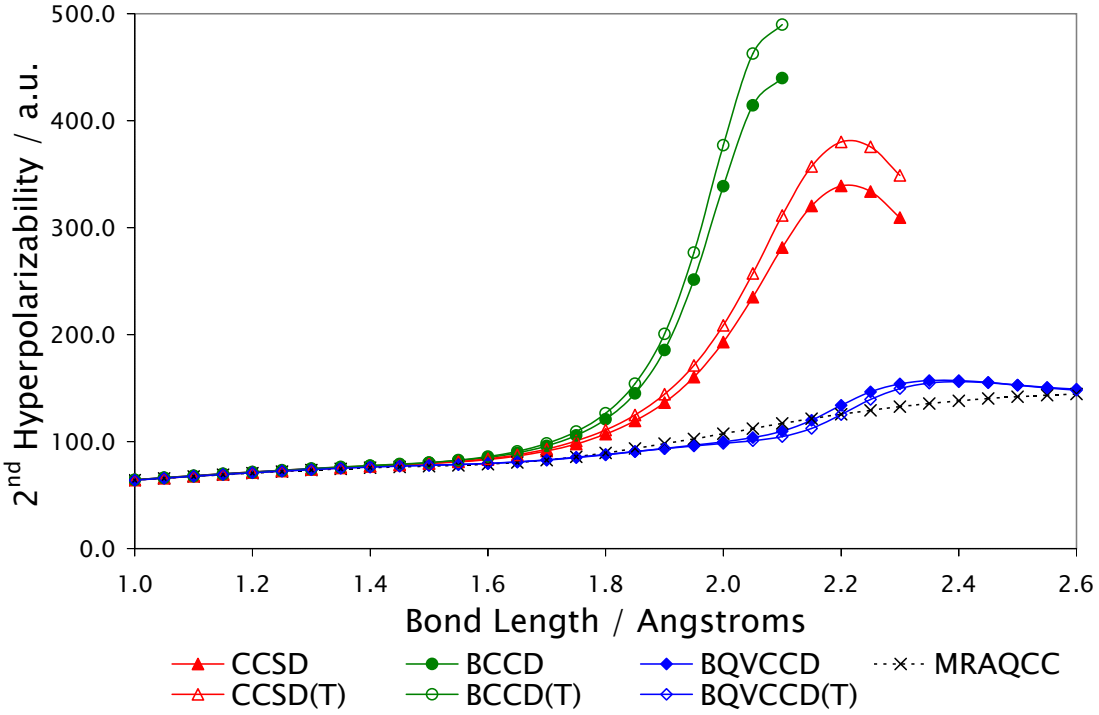


Figure 8.10: Calculated second hyperpolarizabilities for the $D_{\infty h}$ H_6 model with $R_1 = R_2$ and with the aug-cc-pVDZ basis.

8.8 Towards the Metal-Insulator Transition: H_{10}

As a last example, the NLO properties *parallel* to the $D_{\infty h}$ H_{10} model system as a function of the separation between sites are examined. Thus, this example

Table 8.6: Errors relative to MRAQCC for calculated energies of, and polarizabilities perpendicular to the $D_{\infty h}$ H_6 model with the aug-cc-pVDZ basis. Results were obtained from the set of points $\{(R_1, R_2)\}$ where $R_1, R_2 \in \{1.0, 1.2, 1.4, 1.6, 1.8\}$Å.

		CCSD	CCSD(T)	BCCD	BCCD(T)	BQVCCD	BQVCCD(T)
E	$\bar{\epsilon}$	3.03×10^{-3}	-4.09×10^{-3}	2.54×10^{-3}	-3.29×10^{-3}	5.42×10^{-3}	1.12×10^{-4}
	$\bar{\lvert\epsilon\rvert}$	3.52×10^{-3}	4.33×10^{-3}	3.74×10^{-3}	3.53×10^{-3}	5.42×10^{-3}	5.57×10^{-4}
	$\sqrt{\bar{\epsilon^2}}$	4.24×10^{-3}	7.70×10^{-3}	4.64×10^{-3}	6.94×10^{-3}	6.56×10^{-3}	6.27×10^{-4}
	$\lceil\epsilon\rceil$	8.74×10^{-3}	4.35×10^{-4}	8.34×10^{-3}	4.15×10^{-4}	1.25×10^{-2}	9.52×10^{-4}
	$\lfloor\epsilon\rfloor$	-4.33×10^{-3}	-2.21×10^{-2}	-1.18×10^{-2}	-2.60×10^{-2}	6.29×10^{-4}	-1.18×10^{-5}
	σ	3.03×10^{-3}	6.65×10^{-3}	3.97×10^{-3}	6.24×10^{-3}	3.77×10^{-3}	6.29×10^{-4}
α	$\bar{\epsilon}$	-1.35×10^{-1}	-2.35×10^{-1}	-1.61×10^{-1}	-2.16×10^{-1}	9.16×10^{-3}	-2.40×10^{-2}
	$\bar{\lvert\epsilon\rvert}$	1.36×10^{-1}	2.35×10^{-1}	1.61×10^{-1}	2.18×10^{-1}	4.20×10^{-2}	2.50×10^{-2}
	$\sqrt{\bar{\epsilon^2}}$	2.68×10^{-1}	4.39×10^{-1}	3.63×10^{-1}	4.96×10^{-1}	6.02×10^{-2}	3.89×10^{-2}
	$\lceil\epsilon\rceil$	7.80×10^{-3}	-5.09×10^{-3}	6.09×10^{-4}	2.41×10^{-2}	1.49×10^{-1}	8.07×10^{-3}
	$\lfloor\epsilon\rfloor$	-9.27×10^{-1}	$-1.36\times10^{+0}$	$-1.56\times10^{+0}$	$-2.08\times10^{+0}$	-1.14×10^{-1}	-9.86×10^{-2}
	σ	2.36×10^{-1}	3.79×10^{-1}	3.32×10^{-1}	4.56×10^{-1}	6.07×10^{-2}	3.13×10^{-2}
γ	$\bar{\epsilon}$	$5.74\times10^{+0}$	$5.82\times10^{+0}$	$6.23\times10^{+0}$	$5.90\times10^{+0}$	9.74×10^{-1}	1.09×10^{-1}
	$\bar{\lvert\epsilon\rvert}$	$5.74\times10^{+0}$	$6.43\times10^{+0}$	$6.24\times10^{+0}$	$6.62\times10^{+0}$	$1.13\times10^{+0}$	7.70×10^{-1}
	$\sqrt{\bar{\epsilon^2}}$	$1.12\times10^{+1}$	$1.32\times10^{+1}$	$1.46\times10^{+1}$	$1.68\times10^{+1}$	$1.52\times10^{+0}$	$1.18\times10^{+0}$
	$\lceil\epsilon\rceil$	$3.99\times10^{+1}$	$4.86\times10^{+1}$	$6.17\times10^{+1}$	$7.16\times10^{+1}$	$3.69\times10^{+0}$	$2.87\times10^{+0}$
	$\lfloor\epsilon\rfloor$	-1.49×10^{-3}	$-6.61\times10^{+0}$	-1.29×10^{-1}	$-4.58\times10^{+0}$	$-1.63\times10^{+0}$	$-3.28\times10^{+0}$
	σ	$9.79\times10^{+0}$	$1.21\times10^{+1}$	$1.35\times10^{+1}$	$1.61\times10^{+1}$	$1.19\times10^{+0}$	$1.20\times10^{+0}$

is used to investigate whether the extremely positive findings for the calculation of perpendicular polarizabilities using the BQVCCD and BQVCCD(T) methods are reflected in parallel polarizabilities also. The H_{10} system, along with other hydrogen chains of similar length, have been investigated previously in the context of metal-insulator transitions, for example in reference [174]. In order to closely reproduce the results of this paper, and in order to make the FCI calculations practical, use has been made of the minimal STO-3G basis set, rendering the calculated polarizabilities of qualitative validity only.

The calculated energies for this system are given in Figure 8.11, and polarizabilities in Figure 8.12, in which it is clear that each of the methods is capable of describing the short bond length region (the region of increasing slope), where the system is thought to be metallic, but the methods based on TCC struggle in the region of decreasing slope of the polarizability, becoming catastrophically incorrect around a bond length of 1.5Å. The BQVCCD and BQVCCD(T) methods, however, are in good agreement with FCI throughout, despite underestimating the value of the polarizability itself. Thus, not only is the BQVCCD ansatz sim-

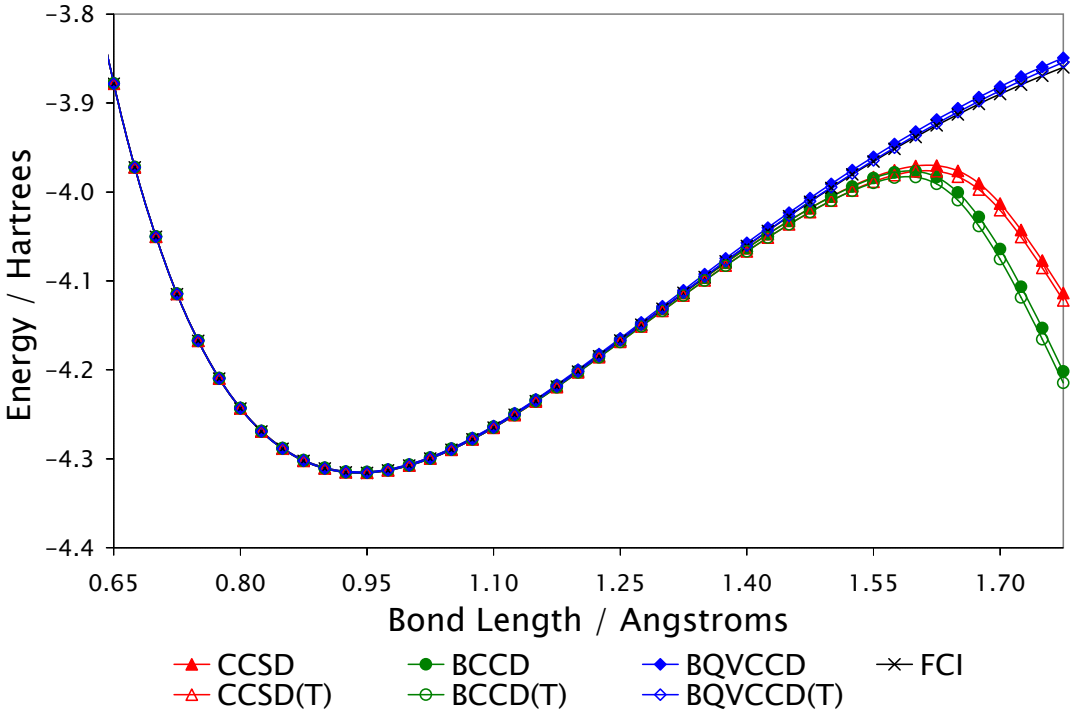


Figure 8.11: Calculated energies of the $D_{\infty h}$ H_{10} model with $R_1 = R_2$ and the STO-3G basis.

ilarly impressive for the evaluation of parallel polarizabilities, but this example illustrates an application of the method to a metal-insulator transition; a problem of widespread interest. For completeness, a plot of the second hyperpolarizability parallel to the longitudinal axis of the molecule in Figure 8.13 is given, which shows similar accuracy.

8.9 Comparison with UHF-CCSD

Finally, as has already been noted, all calculations presented so far have used RHF reference wavefunctions. However, very good results for these model systems can be obtained at extended interatomic distances by the use of a UHF reference wavefunction, which, unlike the RHF wavefunction, is qualitatively correct at dissociation; in RHF theory, the α and β electrons are constrained to occupy the same spatial orbitals. If a molecule dissociates into open-shell fragments, which should be uncharged on physical grounds, this restriction necessarily leads to ionic contamination of the wavefunction as the molecule dissociates[59]. In UHF theory, this constraint is relaxed.

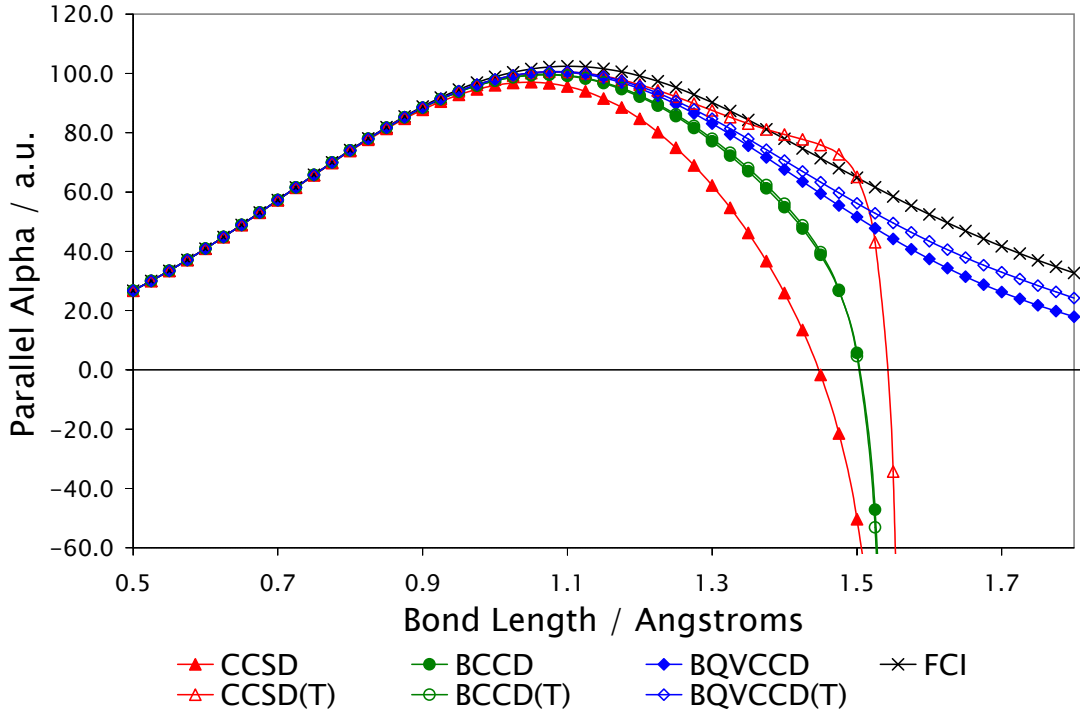


Figure 8.12: Calculated longitudinal polarizabilities of the $D_{\infty h}$ H_{10} model with $R_1 = R_2$ and the STO-3G basis.

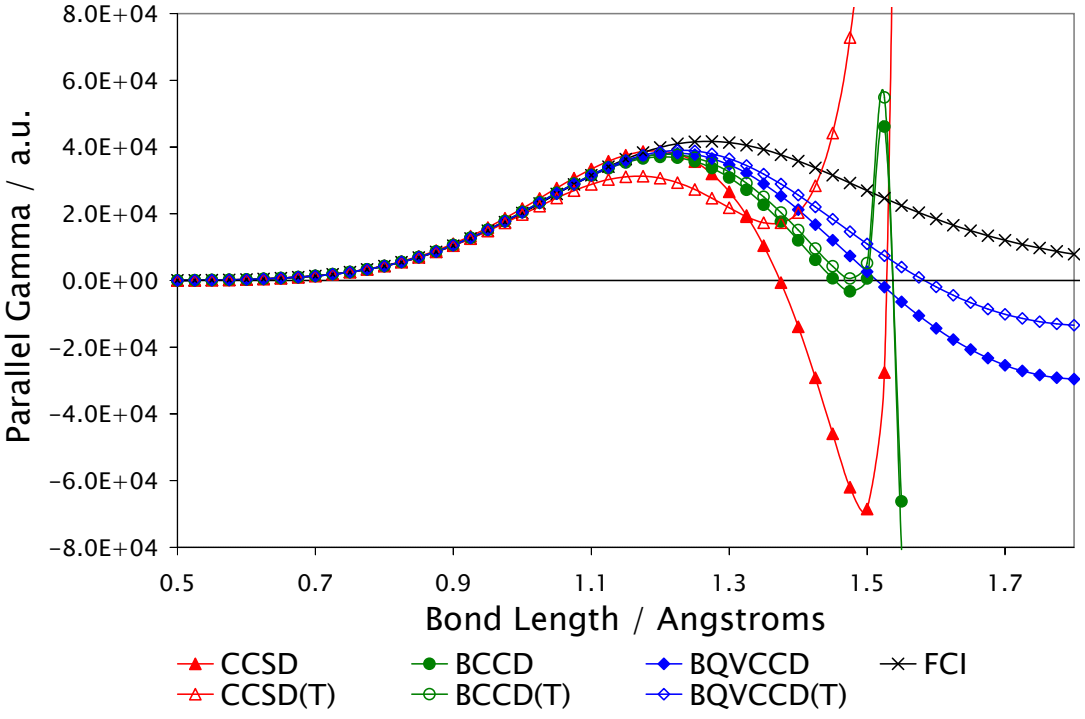


Figure 8.13: Calculated longitudinal second hyperpolarizabilities of the $D_{\infty h}$ H_{10} model with $R_1 = R_2$ and the STO-3G basis.

Thus, it is of interest to explore whether this pseudo-variational method that appears to predict quantitatively accurate potential energy curves for these mod-

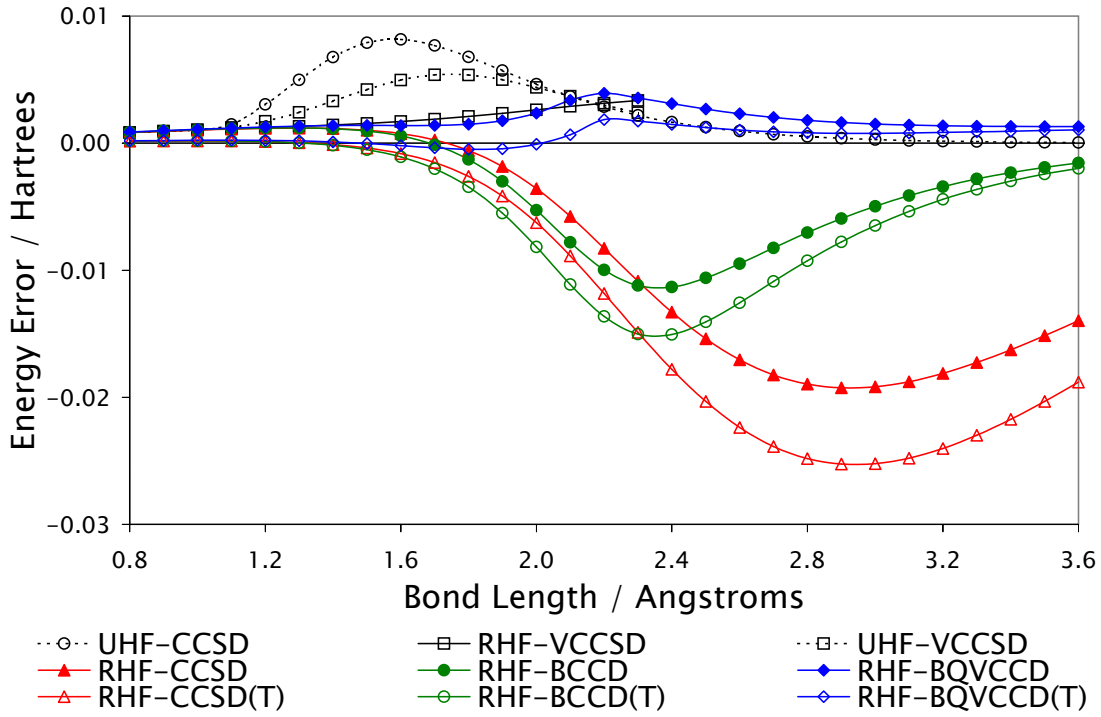


Figure 8.14: Errors relative to FCI for calculated energies of the $D_{\infty h}$ H_4 model with $R_1 = R_2$ and the cc-pVDZ basis.

els produces results comparable to a UHF-based post-Hartree-Fock method. In order to assess this, additional UHF-CCSD calculations were performed on the $D_{\infty h}$ H_4 model. These results, illustrated in Figure 8.14, are quite surprising; as is to be expected, the UHF-CCSD results are in agreement with the RHF-CCSD results at short bond lengths and approach FCI quality at dissociation, but there are large errors in the interval 1.2-2.4Å. This can be ascribed to spin contamination effects[175, 176]. In contrast, an RHF wavefunction is always an eigenfunction of $\hat{S}^2$, so that the RHF-CCSD and RHF-BQVCCD methods yield exact spin eigenstates, and therefore have the advantage over UHF-CCSD that they do not suffer from spin contamination. However, while RHF-CCSD diverges at sufficiently long bond lengths, RHF-BQVCCD displays an accuracy rivalling UHF-CCSD and even exceeding it over regions of the potential energy curve, despite the fact that the reference wavefunction becomes qualitatively wrong.

A more extreme system demonstrating this behaviour can be constructed by arranging six hydrogen atoms uniformly on the circumference of a circle and allowing the radius of the circle to increase; the six-atom equivalent of the H_8 model of the previous chapter. The potential energy curves for this system are given in

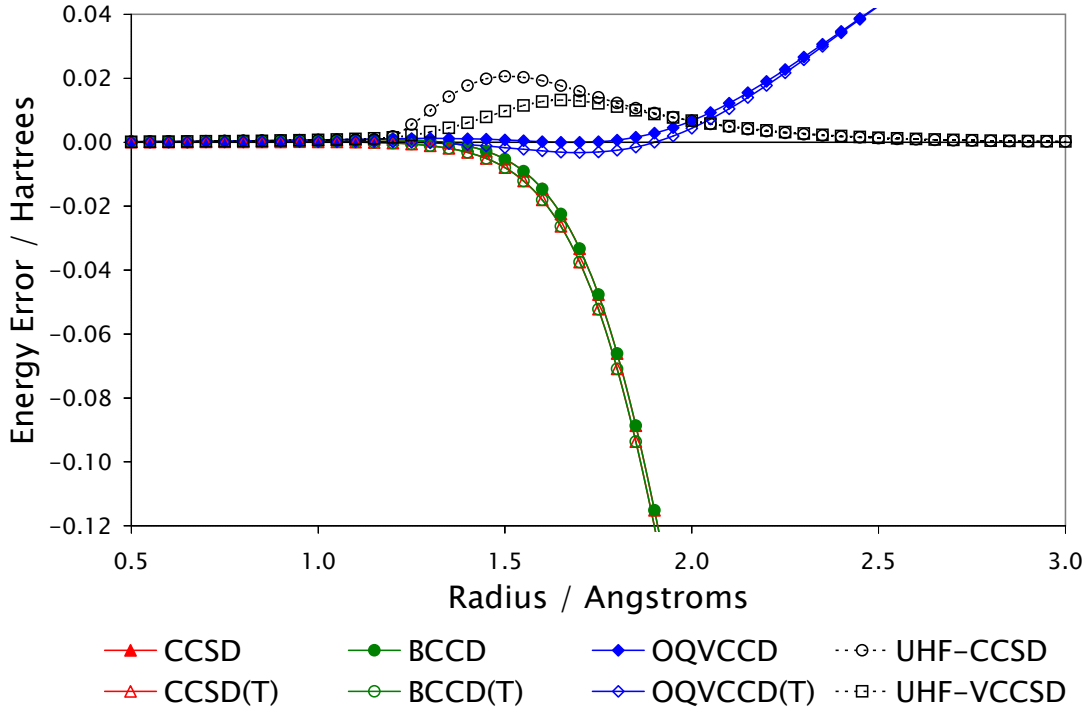


Figure 8.15: Errors relative to FCI for calculated energies for 6 H atoms equally spaced on the circumference of a circle as a function of the radius of the circle, with the STO-3G basis set.

Figure 8.15, with identical conclusions; the UHF-based methods perform better than the RHF-based methods at long interatomic separations, as is to be expected, but there exists an interval of radii for which RHF-CCSD is qualitatively wrong due to the breakdown of the Hartree-Fock approximation, and UHF-CCSD is poor due to spin-contamination, but for which RHF-BQVCCD performs exceptionally well. This suggests the closed-shell-reference QVCC approach to be generally useful and applicable to problems in which an adequate treatment of dissociation is required and a wavefunction that is a spin eigenfunction would also be desirable. In addition, it is clear that UHF-VCC mitigates the effects of spin contamination in comparison to UHF-TCC, as has been noted previously[28]. This gives impetus to the further development of a UHF-QVCC theory.

8.10 Summary

Throughout this chapter, a systematic study of the energies and perpendicular NLO properties of singlet multiradicals, using model hydrogen systems as

prototypical examples, has been carried out. Considering the potential energy curves and associated errors alone, it is clear that RHF-based BQVCCD and BQVCCD(T) significantly out-perform the equivalent RHF-based TCC methods. In the case of the H_4 models, the partially-satisfied variational upper bound property appears to be in effect, allowing the BQVCCD method to remain quantitatively accurate in comparison to FCI, whereas CCSD and BCCD typically fall below FCI and predict poor values. For the cases of H_6 and H_{10}, this poor performance becomes magnified into a readily apparent and catastrophic breakdown of the TCC methods, whereas no corresponding deterioration of BQVCCD or BQVCCD(T) is observed. These results strongly imply that the Quasi-Variational Coupled Cluster ansatz is a suitable black-box method for the highly-correlated treatment of multiradicals, superior to even UHF-based methods at intermediate bond lengths, at which spin contamination can be problematic.

Furthermore, the polarizability study makes apparent the robust nature of the electronic structures calculated by Approximate Variational Coupled Cluster Theories; not only are the calculated energies superior, but high-order properties, such as the second hyperpolarizability, that depend strongly on the quality of the calculated electronic structure are significantly more quantitatively accurate also, sometimes by an order of magnitude, as in the case of H_6. This can be attributed to the approximate enforcement of the variational upper bound property, which appears not just to affect the calculated energies, but, since the bounding of the energy constrains the values of the cluster amplitudes, it has the effect of making the overall description of the electronic structure more robust. In the previous chapter, the stabilization of the perturbative correction for triple excitations, (T), could be explained by this effect, whereas when the same correction is applied using TCC-calculated amplitudes, the correlation energy is catastrophically overestimated. The excellent performance of the BQVCCD method for describing high-order NLO properties can be ascribed to this same effect, and appears to extend not just to perpendicular NLO properties, but to the parallel properties also, such that this family of methods may be of widespread interest.

Chapter 9

Concluding Remarks

This thesis has presented a new family of *ab initio* electronic structure models, entitled *Approximate Variational Coupled Cluster Theories*, which represent new numerical methods for the approximate solution of the electronic Schrödinger equation within the Born-Oppenheimer approximation. These methods are formulated through the minimization of an extensive energy functional that is exact for limiting systems, invariant to rotations in the underlying orbital spaces, and which possesses an intimate relationship with the Variational Coupled Cluster method. As discussed in Chapter 3, starting with the Linked Pair Functional, this family is unlike earlier attempts to construct such approximations that relied on simple truncation of the exponential operator, and instead exploit the internal mathematical structure of VCC; the existence of relationships between VCCD terms guarantees that limiting systems can be treated correctly with only a subset of those terms, albeit an infinite one. These methods are also distinct from improvements upon a TCC framework, since they do not involve the use of the non-Hermitian similarity-transformed Hamiltonian, and do not exceed $\mathcal{O}(o^2v^4)$ complexity.

Chapter 4 gave further technical details of the Linked Pair Functional methodology, including the necessary scheme for the correct minimization of the functional, involving the solution of the eigenproblem for the transformation matrix. In the case of QVCCD, this same procedure is valid, but must be applied to each of the four transformation matrices in turn. The necessity of constructing

positive-definite transformation matrices was additionally discussed, along with the use of optimization of the orbitals or the application of a Brueckner condition for the treatment of single excitations.

In Chapter 5, it was discussed how it is insufficient simply to acquire a representative subset of the terms present in VCCD (even if this subset is correct for limiting systems) if the method does not agree with VCCD well at low orders, and that corrections to the method to improve the low-order approximation of VCCD must be applied. This was extended in Chapter 6, by noting that not only is it necessary to match VCCD at low orders, but in order to avoid the breakdown of the method, the infinite subset must contain balanced contributions through *all* orders. This has resulted in the construction of the Quasi-Variational Coupled Cluster method, the current state of the art, which does not make a distinction between those terms necessary for correct behaviour in the 2-electron limit, and those necessary for the correct low-order behaviour. In fact, this has the effect of restoring hole-particle symmetry, such that the QVCCD functional becomes equivalent to that of CID for a 2-hole system also. It has been conclusively demonstrated that this method inherits pseudo-variational upper bound character from its parent theory, VCCD, and that, in practice, QVCCD predicts potential energy curves analogous to those of VCCD, and distinct from those of CCD when non-dynamic correlation becomes strong.

It was demonstrated in Chapter 7 that the QVCCD (and also AVCCD but not LPFD) method is equivalent to both CCD and VCCD through fourth-order of Møller-Plesset perturbation theory, and (ignoring singles) fully correct to third order, each omitting only those terms containing triple excitations from fourth-order. This allows QVCCD to make use of the same [T] perturbative correction as CCD, which is equivalent to (T) when either optimization of the orbitals or the Brueckner condition is in use, such that the single excitations vanish. The resulting OQVCCD(T) method gives results almost identical with BCCD(T) when Hartree-Fock theory is a good approximation to the true electronic structure. When the Hartree-Fock approximation is invalid, however, and non-dynamic correlation becomes strong, such as when modelling the dissociation of multiply-bonded molecules, not only does OQVCCD significantly out-perform the TCC

methods by predicting physically correct potential energy curves, but the perturbative triples corrections do not significantly deteriorate, unlike when applied in conjunction with CCSD or BCCD, for which a catastrophic overestimation of the correlation energy then occurs; see, for example, Figure 7.12. In particular, the OQVCCD(T) method is capable of predicting a quantitatively accurate potential energy curve for dinitrogen, N_2, challenging the conventional wisdom that a full treatment of quadruple excitations is required to correctly treat this system with a Coupled Cluster methodology based on Restricted Hartree-Fock theory. Furthermore, that a quantitatively accurate complete potential energy curve, from the repulsive domain, through the equilibrium geometry to the dissociation limit, of molecules with electronic structures as complicated as acetylene or dicarbon can be obtained from a strictly single-reference method is simply remarkable.

As discussed more thoroughly in Section 7.4, these surprising results can be ascribed to the approximately-fulfilled variational upper bound property of the OQVCCD method; an upper bound criterion on the energy effectively acts as a constraint on the cluster amplitudes, from which the energy is constructed. In this sense, it can be said that the OQVCCD method allows more robust electronic structures to be predicted than by the TCC methods, and this becomes especially apparent when non-dynamic correlation is strong. The perturbative corrections are themselves constructed from the cluster amplitdes, and therefore benefit indirectly from even the partially-fulfilled upper bound constraint. The presented results also give an important new insight into the nature of the electron correlation problem, because it appears that a Coupled-Cluster-like method that possesses such an approximate upper bound property can adequately capture the essential physics of even severe non-dynamic correlation across a range of systems, and that the remaining dynamic correlation effects then remain legitimately tractable to treatment by perturbation theory.

In Chapter 8, a pilot application of the QVCC methodology (using BQVCCD for direct comparison with BCCD) to the energies, polarizabilities and second hyperpolarizabilities of several model hydrogen systems was presented. The results strongly imply that BQVCCD is an excellent method for the treatment of such systems, and that not only are the predicted energies of these prototypical

multiradicals significantly improved by the upper bound property, but the more robust electronic structures allow the prediction of highly accurate NLO properties also, and this is true not just of perpendicular properties, but of parallel properties, which are of widespread general interest. Since spin contamination is not an issue for the RHF-based methods, the BQVCCD method has been shown also to be highly competitive with UHF-based CCSD methods, which, in contrast to RHF, correctly describe the electron localization that should occur at long bond lengths. It therefore appears that the Quasi-Variational Coupled Cluster method is suitable for the black-box treatment of multiradicals when current single-reference methods perform inadequately, and for which multireference methods are impractical.

The above findings indicate *Approximate Variational Coupled Cluster Theories* to be a promising family of *ab initio* methods, and it is the author's view that significant additional research should be targeted at the further development of these methods. In particular, a few areas of potential improvements and further work can be highlighted.

- At the doubles-only level, QVCCD first differs from VCCD in the 2-electron $\mathcal{O}(T^4)$ terms, or fifth-order in Møller-Plesset theory. This leading-order difference cannot be corrected (at least in the iterative part of the QVCCD calculation) without violating the $\mathcal{O}(o^2v^4)$ complexity criterion, since accounting for the omitted terms would be of $\mathcal{O}(v^6)$ complexity, and are related to the terms captured by the Quadratic Coupled Cluster method[55]. It may be feasible, however, to include such terms in a non-iterative correction to the energy, and to combine it with a more robust perturbative correction for higher excitations, such as (2)[80, 81] or (TQ_f)[51, 177], which also scale with $\mathcal{O}(v^6)$ complexity. Such an *a posteriori* inclusion of these omitted VCCD terms, however, would not further improve the "variational nature" of the QVCCD converged cluster amplitudes, and would therefore lead to no further "surprising" improvement in the efficacy of perturbative corrections, other than that noted for QVCCD alone.

- Despite this limitation, it may be possible to further improve upon the

QVCCD doubles-only approximation to VCCD by accounting for more of the presently omitted higher-order terms, that is, by further enlarging the subsets of VCCD terms captured by QVCCD through all orders. One potential way to do this is to note that, starting first at $\mathcal{O}(T^5)$, some VCCD terms appear to be mixtures of at least two of the terms generated by the $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ transformations. These terms are not captured by QVCCD, since it includes only a linear combination of these transformations. Although these "mixing" terms must cancel for 2 electrons or 2 holes (since QVCCD is exact in those limits), in more general cases they may represent important contributions to the correlation energy, in the same way that the $\mathcal{A}$ and $\mathcal{D}$ terms cancel for 2 electrons, but do not cancel more generally and can therefore become important. Accounting for these omitted terms by allowing for products of the different transformations through some scheme may therefore further improve upon the already excellent QVCCD approximation of VCCD. Whether additional transformations must be introduced to fulfill this goal, and whether it can be achieved without destroying any of the important methodological properties already in place, remains to be seen.

- Since the replacement of single excitations with orbital optimization has an associated increased computational cost, a more thorough investigation of the possibility of explicitly including single excitations should be carried out, with possible further generalization of the method to arbitrary excitation rank. The theory should additionally be extended to treat open-shell systems and excited states, and modern developments in electronic structure theory such as explict correlation[64], which significantly improves the convergence of calculated energies with basis set size, and local correlation[68], which allows the treatment of systems typically far outside the reach of the $\mathcal{O}(N^6)$-scaling Coupled Cluster methods, should be embraced. Furthermore, the potential of extending the *Approximate Variational Coupled Cluster* family of methods to use multireference wavefunctions (as opposed to the single-determinantal reference wavefunctions assumed throughout this thesis) should be investigated; at present, consen-

sus has not been reached on the most appropriate solution to the problem of constructing a multireference Coupled Cluster method that preserves all of the important methodological properties of the single-reference Coupled Cluster paradigm, but since the functional form of the *Approximate Variational Coupled Cluster Theories* discussed in this thesis has more in common with Configuration Interaction and Coupled Pair theories, they may turn out to be feasible alternatives.

With these additional refinements of the theory, this family of methods may become indispensable new tools for a black-box treatment of strongly-correlated or multireference situations, or problems for which CCSD performs inadequately. Further improvements to the quality of the QVCCD approximation to VCCD, in particular, may lead to further surprising improvements in the performance of perturbative corrections for higher excitations, allowing an accurate treatment of problems typically thought to be well outside the reach of single-reference approaches, especially useful when the application of multireference methods becomes problematic.

In conclusion, it is appropriate to devote a final few words to how this thesis should influence the design of future generations of *ab initio* electronic structure models. As evidenced by the fundamental shift that has occurred from the use of Configuration Interaction theory to Traditional Coupled Cluster theory, the contemporary view appears to be that it is acceptable to discard the property of a variational upper bound on the exact ground-state Schrödinger eigenvalue if this allows for the satisfaction of other important properties, such as rigorous extensivity. However, this thesis has conclusively demonstrated even a partially fulfilled variational upper bound property to be extremely effective at extending the range of problems that can be treated within a strictly single-reference Coupled Cluster methodology. The additional enhancement of the efficacy of perturbative corrections for higher excitations that occurs when the physics of a system is then more completely captured by the electronic structure ansatz prompts the re-evaluation of the relative importance of the various methodological features of *ab initio* schemes. The powerful property of a variational upper bound *should not*, and indeed *need not*, be entirely sacrificed.

List of Figures

List of Tables

Bibliography

[1] D. R. Hartree, Proc. Cambridge Phil. Soc. **24**, 89 (1928).

[2] V. Fock, Z. Phys. **61**, 126 (1930).

[3] C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).

[4] N. C. Handy, P. J. Knowles, and K. Somasundram, Theor. Chem. Acc. **68**, 87 (1985).

[5] O. Christiansen, J. Olsen, P. Jørgensen, H. Koch, and P.-Å. Malmqvist, Chem. Phys. Lett. **261**, 369 (1996).

[6] J. Olsen, P. Jørgensen, T. Helgaker, and O. Christiansen, J. Chem. Phys. **112**, 9736 (2000).

[7] P.-O. Löwdin, Phys. Rev. **97**, 1474 (1955).

[8] R. Ahlrichs, P. Scharf, and C. Ehrhardt, J. Chem. Phys. **82**, 890 (1985).

[9] D. P. Chong and S. R. Langhoff, J. Chem. Phys. **84**, 5606 (1986).

[10] R. Ahlrichs, J. Chem. Phys. **48**, 1819 (1968).

[11] W. Meyer, J. Chem. Phys. **58**, 1017 (1973).

[12] C. Kollmar and F. Neese, Molecular Physics **108**, 2449 (2010).

[13] W. Kutzelnigg, Unconventional Aspects of Coupled-Cluster Theory, in *Recent Progress in Coupled Cluster Methods*, edited by P. Čársky, J. Paldus, and J. Pittner, pages 1–61, Springer, 2010.

[14] W. Kutzelnigg and P. von Herigonte, *Electron Correlation at the Dawn of The 21st century*, volume 36 of *Advances in Quantum Chemistry*, Academic Press, 1999.

[15] S. Koch and W. Kutzelnigg, Theor. Chim. Acta **59**, 387 (1981).

[16] R. Ahlrichs, Comput. Phys. Commun. **17**, 31 (1979).

[17] S. R. Langhoff and E. R. Davidson, Int. J. Quantum Chem. **8**, 61 (1974).

[18] M. Nooijen and R. J. LeRoy, J. Mol. Struct.: THEOCHEM **768**, 25 (2006).

[19] F. Coester, Nucl. Phys. **7**, 421 (1958).

[20] F. Coester and H. Kümmel, Nucl. Phys. **17**, 477 (1960).

[21] J. Čížek, J. Chem. Phys. **45**, 4256 (1966).

[22] J. Čížek, Advan. Chem. Phys. **14**, 35 (1969).

[23] J. Čížek and J. Paldus, Int. J. Quantum Chem. **5**, 359 (1971).

[24] R. J. Barlett and G. D. Purvis III, Int. J. Quantum Chem. **14**, 561 (1978).

[25] R. J. Bartlett, Ann. Rev. Phys. Chem. **32**, 359 (1981).

[26] G. D. Purvis III and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).

[27] E. Merzbacher, *Quantum Mechanics*, Wiley, New York, 2nd edition, 1970.

[28] B. Cooper and P. J. Knowles, J. Chem. Phys. **133**, 234102 (2010).

[29] H.-J. Werner and P. J. Knowles, J. Chem. Phys. **89**, 5803 (1988).

[30] P. J. Knowles and H.-J. Werner, Chem. Phys. Lett. **145**, 514 (1988).

[31] R. J. Gdanitz and R. Ahlrichs, Chem. Phys. Lett. **143**, 413 (1988).

[32] P. G. Szalay and R. J. Bartlett, Chem. Phys. Lett. **214**, 481 (1993).

[33] H.-J. Werner and P. J. Knowles, Theor. Chem. Acc. **78**, 175 (1991).

[34] A. Krylov, C. Sherrill, E. Byrd, and M. Head-Gordon, J. Chem. Phys. **109**, 10669 (1998).

[35] J. Parkhill and M. Head-Gordon, J. Chem. Phys. **133**, 124102 (2010).

[36] J. Parkhill and M. Head-Gordon, J. Chem. Phys. **133**, 024103 (2010).

[37] A. Krylov, Chem. Phys. Lett. **338**, 375 (2001).

[38] A. Krylov, Chem. Phys. Lett. **350**, 522 (2001).

[39] M. Nooijen and R. Bartlett, J. Chem. Phys. **107**, 6812 (1997).

[40] K. Sattelmeyer, H. Schaefer, and J. Stanton, Chem. Phys. Lett. **378**, 42 (2003).

[41] P.-D. Fan and P. Piecuch, Adv. Quantum Chem. **51**, 1 (2006).

[42] P. Piecuch, I. S. O. Pimienta, P.-D. Fan, and K. Kowalski, ACS symposium series **958**, 37 (2007).

[43] P. Piecuch, K. Kowalski, P.-D. Fan, and I. S. O. Pimienta, Progress in Theoretical Chemistry and Physics **12**, 119 (2010).

[44] T. Van Voorhis and M. Head-Gordon, J. Chem. Phys. **113**, 8873 (2000).

[45] P.-D. Fan, K. Kowalski, and P. Piecuch, Molec. Phys. **103**, 2191 (2005).

[46] W. Kutzelnigg, J. Chem. Phys. **77**, 3081 (1982).

[47] W. Kutzelnigg and S. Koch, J. Chem. Phys. **79**, 4315 (1983).

[48] W. Kutzelnigg, J. Chem. Phys. **82**, 4166 (1984).

[49] H. F. Schaefer, *Methods of Electronic Structure Theory*, Plenum Press, New York, 1977.

[50] W. Kutzelnigg, Theor. Chem. Acc. **80**, 349 (1991).

[51] R. J. Bartlett and J. Noga, Chem. Phys. Lett. **150**, 29 (1988).

[52] W. Kutzelnigg, Theo. Chim. Acta **80**, 349 (1991).

[53] J. Arponen, Annals of Physics **151**, 311 (1983).

[54] W. Kutzelnigg, Mol. Phys. **94**, 65 (1998).

[55] T. Van Voorhis and M. Head-Gordon, Chem. Phys. Lett. **330**, 585 (2000).

[56] F. A. Evangelista, J. Chem. Phys. **134**, 224102 (2011).

[57] P. Dirac, Proc. R. Soc. A **123**, 714 (1929).

[58] P. Atkins and R. Friedman, *Molecular Quantum Mechanics*, Oxford University Press, 4th edition, 2007.

[59] A. Szabo and N. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Publications, 1996.

[60] H. A. Anton and C. Rorres, *Elementary Linear Algebra*, Wiley, 9th edition, 2005.

[61] M. Boas, *Mathematical Methods in the Physical Sciences*, Wiley, 2nd edition, 1983.

[62] K. Riley, M. Hobson, and S. Bence, *Mathematical Methods for Physics and Engineering*, Cambridge University Press, 2nd edition, 2004.

[63] E. Hylleraas, Z. Phys. **48**, 469 (1928).

[64] W. Klopper, F. R. Manby, S. Ten-No, and E. F. Valeev, Int. Rev. Phys. Chem. **25**, 427 (2006).

[65] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).

[66] R. A. Kendall, T. H. Dunning Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[67] P. G. Szalay, M. Nooijen, and R. J. Bartlett, J. Chem. Phys. **103**, 281 (1995).

[68] C. Hampel and H.-J. Werner, J. Chem. Phys. **104**, 6286 (1996).

[69] A. Einstein, Ann. der Physik **354**, 769 (1916).

[70] J. A. Pople, M. Head-Gordon, and K. Raghavachari, J. Chem. Phys. **87**, 5968 (1987).

[71] J. Hrušák, S. Ten-no, and S. Iwata, J. Chem. Phys. **106**, 7185 (1997).

[72] C. Kollmar, J. Chem. Phys. **125**, 084108 (2006).

[73] C. Kollmar and F. Neese, J. Chem. Phys. **135**, 084102 (2011).

[74] L. M. J. Huntington and M. Nooijen, J. Chem. Phys. **133**, 184109 (2010).

[75] D. A. Mazziotti, Phys. Rev. Lett. **93**, 213001 (2004).

[76] D. A. Mazziotti, Acc. Chem. Res. **39**, 207 (2006).

[77] A. DePrince III, E. Kamarchik, and D. Mazziotti, J. Chem. Phys. **128**, 234103 (2008).

[78] A. Sand, C. Schwerdtfeger, and D. Mazziotti, J. Chem. Phys. **136**, 034112 (2012).

[79] T. Crawford and H. Schaefer, Rev. Comp. Chem. **14**, 33 (2007).

[80] S. R. Gwaltney, C. D. Sherrill, and M. Head-Gordon, J. Chem. Phys. **113**, 3548 (2000).

[81] S. R. Gwaltney and M. Head-Gordon, J. Chem. Phys. **115**, 2014 (2001).

[82] S. Kucharski and R. Bartlett, J. Chem. Phys. **108**, 5243 (1998).

[83] A. Taube and R. Bartlett, J. Chem. Phys. **128**, 044110 (2008).

[84] A. Taube and R. Bartlett, J. Chem. Phys. **128**, 044111 (2008).

[85] M. Musial and R. J. Bartlett, J. Chem. Phys. **133**, 104102 (2010).

[86] H.-J. Werner et al., Molpro, version 2010.2, a package of ab initio programs, 2011, see http://www.molpro.net.

[87] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, WIREs Comput. Mol. Sci. **2**, 242 (2012).

[88] R. Bartlett and M. Musial, Rev. Mod. Phys. **79**, 291 (2007).

[89] P. J. Knowles and B. Cooper, J. Chem. Phys. **133**, 224106 (2010).

[90] J. M. Howie, *Real Analysis*, Springer, 1st edition, 2001.

[91] J. M. Howie, *Complex Analysis*, Springer, 1st edition, 2003.

[92] W. Meyer, J. Chem. Phys. **58**, 1017 (1973).

[93] W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. **110**, 2800 (1998).

[94] J. B. Robinson and P. J. Knowles, J. Chem. Phys. **135**, 044113 (2011).

[95] J. B. Robinson and P. J. Knowles, J. Chem. Phys. **136**, 054114 (2012).

[96] N. C. Handy, J. A. Pople, M. Head-Gordon, K. Raghavachari, and G. W. Trucks, Chem. Phys. Lett. **164**, 185 (1989).

[97] K. A. Brueckner, Phys. Rev. **96**, 508 (1954).

[98] R. K. Nesbet, Phys. Rev. **109**, 1632 (1958).

[99] D. J. Thouless, *The Quantum Mechanics of Many-Body Systems*, Academic, New York, 1961.

[100] C. Kollmar and A. Heßelmann, Theor. Chem. Acc. **127**, 311 (2009).

[101] L. Adamowicz, W. D. Laidig, and R. J. Bartlett, Int. J. Quantum Chem. Symp. **18**, 245 (1984).

[102] C. Hampel, K. Peterson, and H.-J. Werner, Chem. Phys. Lett. **190**, 1 (1992).

[103] T. D. Crawford and J. F. Stanton, J. Chem. Phys. **112**, 7873 (2000).

[104] J. B. Robinson and P. J. Knowles, Phys. Chem. Chem. Phys. **14**, 6729 (2012).

[105] J. B. Robinson and P. J. Knowles, J. Chem. Theory Comput. (2012), *"Benchmark Quasi-Variational Coupled Cluster Calculations of Multiple Bond Breaking"*, in press.

[106] M. Urban, J. Noga, S. J. Cole, and R. J. Bartlett, J. Chem. Phys. **83**, 4041 (1985).

[107] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).

[108] J. F. Stanton, Chem. Phys. Lett. **281**, 130 (1997).

[109] M. J. O. Deegan and P. J. Knowles, Chem. Phys. Lett. **227**, 321 (1994).

[110] W. D. Laidig, P. Saxe, and R. J. Bartlett, J. Chem. Phys. **86**, 887 (1987).

[111] S. R. Gwaltney and M. Head-Gordon, Chem. Phys. Lett. **323**, 21 (2000).

[112] P. Piecuch and M. Włoch, J. Chem. Phys. **123**, 224105 (2005).

[113] S. R. Langhoff and E. R. Davidson, Int. J. Quantum Chem. **8**, 61 (1974).

[114] M. W. Schmidt et al., J. Comput. Chem. **14**, 1347 (1993).

[115] F. Neese, Orca - an ab initio, density functional and semiempirical program package, version 2.6, 2008.

[116] K. P. Huber and G. Herzberg, *Constants of Diatomic Molecules*, Van Nostrand Reinhold, 1979.

[117] Y. Ge, M. Gordon, and P. Piecuch, J. Chem. Phys. **127**, 174106 (2007).

[118] M. Kállay and P. R. Surján, J. Chem. Phys. **115**, 2945 (2001), see also http://www.mrcc.hu.

[119] J. B. Robinson and P. J. Knowles, J. Chem. Phys. (2012), *"Application of the Quasi-Variational Coupled Cluster Method to the Non-Linear Optical Properties of Model Hydrogen Systems"*, in press.

[120] H. Nalwa and S. Miyata, *Nonlinear Optics of Organic Molecules and Polymers*, CRC Press, 1st edition, 1997.

[121] B. H. Cumpston et al., Nature **398**, 51 (1999).

[122] S. Kawata, H.-B. Sun, T. Tanaka, and K. Takada, Nature **412**, 697 (2001).

[123] M. Albota et al., Science **281**, 1653 (1998).

[124] W. Zhou et al., Science **296**, 1106 (2002).

[125] W. R. Dichtel et al., J. Am. Chem. Soc. **126**, 5380 (2004).

[126] F. Terenziani, C. Katan, E. Badaeva, S. Tretiak, and M. Blanchard-Desce, Adv. Mater. **20**, 4641 (2008).

[127] Z. Li et al., Chem. Phys. Lett. **441**, 123 (2007).

[128] W. Zhao, J. Phys. Chem. Lett. **2**, 482 (2011).

[129] M. Samoc et al., Inorg. Chem. **47**, 9946 (2008).

[130] R. Zalesny et al., Phys. Chem. Chem. Phys. **12**, 373 (2010).

[131] S. Ohira et al., J. Am. Chem. Soc. **131**, 6099 (2009).

[132] H. Reis et al., J. Comp. Chem. **32**, 908 (2011).

[133] R. Orlando et al., J. Phys. Chem. A **115**, 12631 (2011).

[134] S. R. Marder et al., Science **261**, 186 (1993).

[135] F. Meyers, S. R. Marder, B. M. Pierce, and J. L. Bredas, J. Am. Chem. Soc. **116**, 10703 (1994).

[136] J. L. Bredas, C. Adant, P. Tackx, A. Persoons, and B. M. Pierce, Chem. Rev. **94**, 243 (1994).

[137] R. R. Tykwinski et al., J. Phys. Chem. B **102**, 4451 (1998).

[138] B. Champagne and E. A. Perpete, Int. J. Quantum Chem. **75**, 441 (1999).

[139] B. Champagne et al., J. Phys. Chem. A **104**, 4755 (2000).

[140] S. Borini, P. A. Limacher, and H. P. Luthi, J. Chem. Phys. **131**, 124105 (2009).

[141] M. de Wergifosse and B. Champagne, J. Chem. Phys. **134**, 074113 (2011).

[142] M. Nakano, T. Minami, R. Kishi, Y. Shigeta, and B. Champagne, J. Chem. Phys. **136**, 024315 (2012).

[143] M. Nakano et al., J. Phys. Chem. A **109**, 885 (2005).

[144] M. Nakano et al., Chem. Phys. Lett. **467**, 120 (2008).

[145] H. Nagai et al., Chem. Phys. Lett. **489**, 212 (2010).

[146] K. Kamada et al., Angewandte Chemie International Edition **46**, 3544 (2007).

[147] H. Kishida, K. Hibino, A. Nakamura, D. Kato, and J. Abe, Thin Solid Films **519**, 1028 (2010).

[148] R. J. Bartlett and G. D. Purvis III, Phys. Rev. A **20**, 1313 (1979).

[149] P. Limacher, L. Qingxu, and H. Lüthi, J. Chem. Phys. **135**, 014111 (2011).

[150] H. Sekino, Y. Maeda, M. Kamiya, and K. Hirao, J. Chem. Phys. **126**, 014107 (2007).

[151] J.-W. Song, M. Watson, H. Sekino, and K. Hirao, J. Chem. Phys. **129**, 024117 (2008).

[152] B. Kirtman et al., J. Chem. Phys. **128**, 114108 (2008).

[153] T. Toto, J. Toto, C. de Melo, M. Hasan, and B. Kirtman, Chem. Phys. Lett. **244**, 59 (2009).

[154] J.-W. Song, M. Watson, H. Sekino, and K. Hirao, Int. J. Quantum Chem. **109**, 2012 (2009).

[155] B. Champagne and B. Kirtman, Int. J. Quantum Chem. **109**, 3103 (2009).

[156] P. A. Limacher, K. V. Mikkelsen, and H. P. Luthi, J. Chem. Phys. **130**, 194114 (2009).

[157] H. Sekino and R. J. Bartlett, J. Chem. Phys. **98**, 3022 (1993).

[158] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[159] W. Kohn and L. Sham, Phys. Rev. **140**, A1133 (1965).

[160] B. Champagne et al., J. Chem. Phys. **109**, 10489 (1998).

[161] S. J. A. van Gisbergen et al., Phys. Rev. Lett. **83**, 694 (1999).

[162] P. Mori-Sanchez, Q. Wu, and W. Yang, J. Chem. Phys. **119**, 11001 (2003).

[163] M. Kamiya, H. Sekino, T. Tsuneda, and K. Hirao, J. Chem. Phys. **122**, 234111 (2005).

[164] B. Champagne and D. Mosley, J. Chem. Phys. **105**, 3592 (1996).

[165] M. Nakano et al., J. Chem. Phys. **125**, 074113 (2006).

[166] M. Nakano, H. Nagao, and K. Yamaguchi, Phys. Rev. A **55**, 1503 (1997).

[167] M. Nakano et al., Chem. Phys. Lett. **432**, 473 (2006).

[168] M. Nakano, S. Yamada, S. Kiribayashi, and K. Yamaguchi, Int. J. Quantum Chem. **70**, 269 (1998).

[169] S. Wouters, P. A. Limacher, D. V. Neck, and P. W. Ayers, J. Chem. Phys. **136**, 134110 (2012).

[170] B. Champagne, D. H. Mosley, M. Vračko, and J.-M. André, Phys. Rev. A **52**, 178 (1995).

[171] B. Champagne, D. H. Mosley, M. Vračko, and J.-M. André, Phys. Rev. A **52**, 1039 (1995).

[172] J. L. Toto, T. T. Toto, and C. P. de Melo, Chem. Phys. Lett. **245**, 660 (1995).

[173] Q. Li, L. Chen, Q. Li, and Z. Shuai, Chem. Phys. Lett. **457**, 276 (2008).

[174] G. Bendazzoli, S. Evangelisti, and A. Monari, Int. J. Quantum Chem. **111**, 3416 (2011).

[175] R. H. Nobes, J. A. Pople, L. Radom, N. C. Handy, and P. J. Knowles, Chem. Phys. Lett. **138**, 481 (1987).

[176] P. J. Knowles and N. C. Handy, J. Phys. Chem. **92**, 3097 (1988).

[177] S. Kucharski and R. J. Bartlett, J. Chem. Phys. **108**, 9221 (1998).

Declaration

- This thesis is submitted to Cardiff University in partial fulfilment of the requirements for the degree of Doctor of Philosophy (Ph.D.).
- This work has not been previously submitted in substance at this or any other university or place of learning, nor is it being submitted concurrently in candidature for any other degree or award.
- This thesis is the result of the author's own independent work and investigation, except where otherwise stated. Other sources are acknowledged by explicit references.
- This thesis may, upon acceptance, be made available for photocopying and for inter-library loan, and the title and abstract may be made available to outside organizations.

Candidate Signature: (James B. Robinson)